Range-based Parameter Estimation in Diffusion Models

Statistical Concepts and Analytical Foundations

DISSERTATION

zur Erlangung des akademischen Grades

Dr. rer. nat.
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Dipl.–Math. Hartmuth Henkel
geb. am 27.08.1981 in Mannheim

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Peter Frensch

Gutachter:
1. Prof. Dr. Markus Reiß
2. Prof. Dr. Dirk Becherer
3. Prof. Dr. Emmanuel Gobet

eingereicht am: 29. März 2010
Tag der mündlichen Prüfung: 24. September 2010

Abstract

We study the behavior of the maximum, the minimum and the terminal value of time–homogeneous one–dimensional diffusions on finite time intervals. To begin with, we prove an existence result for the joint density by means of Malliavin calculus. Moreover, we derive expansions for the joint moments of the triplet (H,L,X) at time Delta w.r.t. Delta. Here, X stands for the underlying diffusion whereas H and L denote its running maximum and its running minimum, respectively. In a first approach that entirely relies on elementary estimates, such as Doob's inequality and Cauchy–Schwarz' inequality, we derive an expansion w.r.t. the square root of the time parameter Delta including powers of 2. A more sophisticated ansatz uses partial differential equation techniques to determine an expansion of the one–barrier hitting time probability for pinned diffusions. For an expansion of the transition density of diffusions is known, one obtains an overall expansion of the joint probability of (H,X) w.r.t. Delta.

The developed distributional properties enable us to establish a theory for martingale estimating functions constructed from range–based data in a parameterized diffusion model. A small–Delta–optimality approach, that uses the approximated moments, yields a simplification of the relatively complicated estimating procedure and we obtain asymptotic optimality results when the sampling frequency Delta tends to 0. When it comes to estimating the drift coefficient the range–based method is not superior to the method relying on equidistant observations of the underlying diffusion alone. However, there is an enormous gain in efficiency at the estimation for the diffusion coefficient. Incorporating the maximum and the minimum into the analysis significantly lowers the asymptotic variance of the estimators for the parameter in this scenario.

Keywords: Range in diffusion models, Range–based parameter estimation, Martingale estimating functions, Small–Delta–Optimality

Zusammenfassung

Wir studieren das Verhalten des Maximums, des Minimums und des Endwerts zeithomogener eindimensionaler Diffusionen auf endlichen Zeitintervallen. Zuerst beweisen wir mit Hilfe des Malliavin–Kalküls ein Existenzresultat für die gemeinsamen Dichten. Außerdem leiten wir Entwicklungen der gemeinsamen Momente des Tripels (H,L,X) zur Zeit Delta bzgl. Delta her. Dabei steht X für die zugrundeliegende Diffusion, und H und L bezeichnen ihr fortlaufendes Maximum bzw. Minimum. Ein erster Ansatz, der vollständig auf den elementaren Abschätzungen der Doob'schen und der Cauchy–Schwarz'schen Ungleichung beruht, liefert eine Entwicklung bis zur Ordnung 2 bzgl. der Wurzel der Zeitvariablen Delta. Ein komplexerer Ansatz benutzt Partielle–Differentialgleichungstechniken, um eine Entwicklung der einseitigen Austrittswahrscheinlichkeit für gepinnte Diffusionen zu bestimmen. Da eine Entwicklung der Übergangsdichten von Diffusionen bekannt ist, erhält man eine vollständige Entwicklung der gemeinsamen Wahrscheinlichkeit von (H,X) bzgl. Delta.

Die entwickelten Verteilungseigenschaften erlauben es uns eine Theorie für Martingalschätzfunktionen, die aus wertebereich–basierten Daten konstruiert werden, in einem parameterisierten Diffusionsmodell herzuleiten. Ein small–Delta–Optimalitätsansatz, der die approximierten Momente benutzt, liefert eine Vereinfachung der vergleichsweise komplizierten Schätzprozedur und wir erhalten asymptotische Optimalitätsresultate für gegen 0 gehende Sampling–Frequenz. Beim Schätzen des Drift–Koeffizienten ist der wertebereich–basierte Ansatz der Methode, die auf Equidistanten Beobachtungen der Diffusion beruht, nicht überlegen. Der Effizienzgewinn im Fall des Schätzens des Diffusionskoeffizienten ist hingegen enorm. Die Maxima und Minima in die Analyse miteinzubeziehen senkt die Varianz des Schätzers für den Parameter in diesem Szenario erheblich.

Schlagworte: Wertebereich in Diffusionsmodellen, Wertebereich–basierte Parameterschätzung, Martingal Schätzfunktionen, Small–Delta–Optimalität

Detailed Abstract

We consider a process X, defined by the stochastic differential equation

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = x,\ t \geq 0.$$

Here B denotes the standard Brownian motion of $\mathbb{R}$ and the coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ are supposed to be sufficiently smooth functions that are parameterized by a parameter $\theta \in \Theta \subset \mathbb{R}$. In the present thesis, we establish inference methods for the parameter θ that are based on the observation of the triplet (H, L, X), where the processes H and L are formally defined by

$$H_t = \sup_{0\leq s\leq t} X_s, \quad \text{and} \quad L_t = \inf_{0\leq s\leq t} X_s.$$

As a very first step, for $t > 0$, we prove an existence result for the joint density of (H_t, L_t, X_t), conditional on $X_0 = x$, by means of Malliavian calculus. In addition, possibilities are presented to calculate these densities - at least theoretically. These results put us into a position to derive a generalized theory for so-called *martingale estimating functions.* Briefly speaking, for a fixed sampling frequency Δ, these estimating functions are constructed from the observations $(H_{\Delta i}, L_{\Delta i}, X_{\Delta i})$ on disjoint intervals $(\Delta(i-1), \Delta i]$, $i = 1, \ldots, n$. In this context, $H_{\Delta i} = \sup_{\Delta(i-1)\leq s\leq \Delta i} X_s$ and $L_{\Delta i} = \inf_{\Delta(i-1)\leq s\leq \Delta i} X_s$. We prove consistency and asymptotic normality of the resulting estimators as the number of observations n tends to ∞. Moreover, we introduce optimality criteria and we scrutinize on which conditions the generalized martingale estimating functions are optimal.

The existence result for the joint density and the results concerning martingale estimating functions are highly theoretical because the joint densities or the joint distributions cannot be calculated explicitly in general. This is the reason why we also focus on the search for alternative inference methods. A canonical way to simplify the estimating procedure is to approximate the aforementioned martingale estimating functions by their first or second order approximations. Therefore we try to find an expansion of the expression $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$. Here, $g : \mathbb{R}^3 \to \mathbb{R}$ can be any sufficiently smooth function that does not grow too fast. A first approach, that relies entirely on elementary estimates, yields an expansion including powers of 2, that is the highest order appearing in this expansion is $\sqrt{t}^2 = t$. This result already suffices to state a *small-Δ-optimality property.* Concretely, this property concerns approximately optimal estimating functions, that are constructed from a fixed number n of observations, when the sampling frequency Δ tends to 0. But, as a simulation shows, the resulting small-Δ-optimal estimators do not perform very well for relatively large observation intervals. This is clearly due to the error induced by the approximation. In order to determine more accurate estimators, a higher order expansion of the quantity $\mathbb{E}_x[g(H_t, L_t, X_t)]$ is required.

A partial differential equation approach yields an overall expansion of the hitting time probability

$$\mathbb{P}_x[\tau_h \leq t | X_t = y]$$

for a class of pinned diffusions, where $\tau_h = \inf\{t > 0 | X_t \geq h\}$. This result can be used to calculate an expansion of $\mathbb{E}_x[g(H_t, X_t)]$ with respect to $\sqrt{t}$. To exemplify the improvements involved, some of the higher order terms are calculated explicitly.

Detaillierte Zusammenfassung

Wir betrachten einen Prozess X, der durch die Stochastische Differentialgleichung

$$dX_t = b(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = x,\ t \geq 0,$$

definiert ist. Dabei bezeichnet B die gewöhnliche Brown'sche Bewegung auf $\mathbb{R}$ und die Koeffizienten $\mu : \mathbb{R} \to \mathbb{R}$, und $\sigma : \mathbb{R} \to \mathbb{R}_+$ sollen hinreichend glatte Funktionen sein, die durch $\theta \in \Theta \subset \mathbb{R}$ parametrisiert sind. In der vorliegenden Dissertation leiten wir Schätzmethoden für den Parameter θ her, die auf der Beobachtung des Vektors (H, L, X) beruhen, wobei die Prozesse H und L formell durch die folgenden Ausdrücke definiert sind:

$$H_t = \sup_{0 \leq s \leq t} X_s, \quad \text{bzw.} \quad L_t = \inf_{0 \leq s \leq t} X_s.$$

Als allerersten Schritt beweisen wir mit Hilfe des Malliavin-Kalküls ein Existenzresultat für die gemeinsame Dichte von (H, L, X). Zusätzlich zeigen wir Wege auf, die Dichten - zumindest theoretisch - zu berechnen. Diese Resultate versetzen uns in die Lage, eine verallgemeinerte Theorie für sogenannte *Martingal Schätzfunktionen* herzuleiten. Für eine feste Beobachtungsfrequenz Δ beruhen diese Schätzfunktionen auf den Beobachtungen $(H_{\Delta i}, L_{\Delta i}, X_{\Delta i})$, für disjunkte Intervalle $(\Delta(i-1), \Delta i]$, $i = 1, \ldots, n$. In diesem Zusammenhang bedeuten $H_{\Delta i} = \sup_{\Delta(i-1) \leq s \leq \Delta i} X_s$ und $L_{\Delta i} = \inf_{\Delta(i-1) \leq s \leq \Delta i} X_s$. Wir beweisen die Konsistenz und die asymptotische Normalität der resultierenden Schätzer, wenn die Anzahl der Beobachtungen n gegen ∞ geht. Außerdem führen wir Optimalitätskriterien ein, und wir untersuchen, unter welchen Bedingungen unsere verallgemeinerten Martingal Schätzfunktionen optimal sind.

Das Existenzresultat für die gemeinsamen Dichten und die Resultate für Martingal Schätzfunktionen sind hoch theoretischer Natur, da die gemeinsamen Dichten im Allgemeinen nicht explizit berechnet werden können. Deshalb konzentrieren wir uns ebenso auf die Suche nach alternativen Schätzverfahren. Ein kanonischer Weg, die statischen Methoden zu vereinfachen, besteht darin, die eben genannten Martingal Schätzfunktionen durch ihre Approximation erster oder zweiter Ordnung anzunähern. Aus diesem Grund versuchen wir, den Ausdruck $\mathbb{E}_x[g(H_t, L_t, X_t)]$ bezüglich $\sqrt{t}$ zu entwickeln. Dabei ist $g : \mathbb{R}^3 \to \mathbb{R}$ eine hinreichend glatte Funktion, die nicht zu stark wächst. Ein erster Ansatz, der voll und ganz auf elementaren Abschätzungen beruht, liefert eine Entwicklung zweiter Ordnung. Dieses Resultat reicht bereits aus, um eine *Small-Δ-Optimalitätseigenschaft* herzuleiten. Konkret betrifft diese Eigenschaft annähernd optimale Schätzfunktionen, die mit Hilfe einer endlichen Zahl von Beobachtungen n konstruiert sind, wenn die Beobachtungsfrequenz Δ gegen 0 geht. Allerdings zeigt eine Simulationsstudie, dass die small-Δ-optimalen Schätzer für relativ große Beobachtungsintervalle keine besonders guten Ergebnisse liefern. Das ist sicherlich auf den Approximationsfehler zurückzuführen. Um akkuratere Schätzer zu erhalten, benötigt man höhere Entwicklungen von $\mathbb{E}_x[g(H_t, L_t, X_t)]$.

Ein Ansatz für Partielle Differentialgleichungen liefert eine vollständige Entwicklung des Terms

$$\mathbb{P}_x[\tau_h \leq t | X_t = y],$$

für eine Klasse gepinnter Diffusionen, wobei $\tau_h = \inf\{t > 0 | X_t \geq h\}$. Dieses Resultat kann dazu verwendet werden, $\mathbb{E}_x[g(H_t, X_t)]$ bezüglich $\sqrt{t}$ zu entwickeln. Um die Nützlichkeit unserer Ergebnisse zu veranschaulichen, berechnen wir einige Terme höherer Ordnung explizit.

Contents

1 Introduction

Topic

Volatility estimation for financial assets is an issue whose importance is constantly increasing. In recent years computational power and storage possibilities for digital data have significantly improved. This technological progress explains why estimation methods that make use of high-frequency financial data have become more and more popular. A major approach concerns the non-parametric determination of the integrated volatility. It is common practice to estimate this quantity from the sum of the frequently sampled squared returns. Although this is justified on the assumption of a continuous stochastic model in an idealized world, it runs into the challenge from market microstructure in real applications. The usual mechanism for dealing with this problem is to create different reasonable subsample estimators in a first step and then to reassemble these estimators in a second step. Many authors have studied realized and integrated volatilities in the presence of market microstructure noise. We deem the articles of Zhang et al. [74] and Zhang [73] to be good references for a first overview.

Beside the aforementioned microstructure effects, financial data that has been sampled at high-frequency also carries other undesired information. During lunch breaks trading activities significantly decrease. On the contrary, trading activities at stock markets all over the world experience a steep increase the moment Wall Street opens trading. Those artefacts blur the estimates based on high-frequency observations. As for the case of microstructure noise, a natural way to deal with this problem is to reduce the sampling frequency. This can be carried to the point where only one or two observations within a medium-term period are used to estimate the volatility. For example, the opening and the closing prices of a trading day alone could be used. Other canonical candidates are the maximum and the minimum values of a trading day. The range of an asset contains a lot of information about the volatility and it reflects, in some sense, the activities taking place during a period. Sampling at a very low rate is robust against short-term effects, but of course, a considerable amount of data remains unused.

Ideally, one would take a combination of the two extreme approaches into account. A reasonable compromise could be to consider a sampling frequency of a few minutes, and to use the starting point and the end point as well as the range of the asset on the observation interval to construct an estimator for the volatility. But how exactly can the maxima and the minima be incorporated into the analysis? The aim of the present thesis is to depict methods to handle range-based data for general diffusion models and to construct reasonable estimators. Note that recent research results suggest that strictly continuous models are not appropriate for modeling stock markets. Lévy processes with finite jump activity seem to be a better choice. See e.g. Jacod and Aït-Sahalia [39]. Nevertheless, our results can be considered as a first step towards a better understanding of the role the range plays in stock market models and of the impact it has on the analysis of the volatility.

Theoretical foundations

First, let us consider a one-dimensional Brownian motion with drift defined by

$$X_t = \mu t + \sigma B_t, \quad t \geq 0, \tag{1.1}$$

where $(B_t,\, t \geq 0)$ denotes the standard Brownian motion of $\mathbb{R}$ and where the coefficients $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ are constants. The distributional properties of the process X, its running maximum $H_t = \sup_{0 \leq s \leq t} X_s$ and its running minimum $L_t = \inf_{0 \leq s \leq t} X_s$, can be used to construct (range-based) estimators for the diffusion coefficient σ. Let the sampling frequency $\Delta > 0$ be fixed and assume that we observe $X_{\Delta i}$ for $i = 0, \ldots, n$. Moreover, let us suppose that we are given the observations $H_{\Delta i} = \sup_{\Delta(i-1) \leq s \leq \Delta i} X_s$ and $L_{\Delta i} = \inf_{\Delta(i-1) \leq s \leq \Delta i} X_s$, $i = 1, \ldots, n$. That means we are given the maxima, the minima and the terminal values of the process X during the observation intervals $(\Delta(i-1), \Delta i]$ for $i = 1, \ldots, n$. Of course, the most conventional estimator of σ^2 is the estimator

$$\hat{\sigma}^2_{close} = \frac{1}{n\Delta} \sum_{i=1}^{n} (X_{\Delta i} - X_{\Delta(i-1)} - \mu t)^2. \tag{1.2}$$

But also the *Parkinson estimator*

$$\hat{\sigma}^2_{Parkinson} = \frac{1}{n\Delta \log 16} \sum_{i=1}^{n} (H_{\Delta i} - L_{\Delta i})^2, \tag{1.3}$$

see [51], is a reasonable estimator for σ^2. Both estimators rely on very elementary distributional properties of the Brownian motion with drift X and its range $H - L$. In the case of (1.3) the justification for the term "range-based" is obvious. By slightly abusing this term, we sometimes also refer to other estimators as range-based estimators if they incorporate the running maximum or the running minimum in some way. In this sense, a more sophisticated range-based estimator is, for example, given by

$$\hat{\sigma}^2_{RS} = \frac{1}{n\Delta} \sum_{i=1}^{n} (H_{\Delta i} - X_{\Delta i})(H_{\Delta i} - X_{\Delta(i-1)}) + (L_{\Delta i} - X_{\Delta i})(L_{\Delta i} - X_{\Delta(i-1)}). \tag{1.4}$$

This particular estimator was found by Rogers and Satchell, see [61]. Another example is the so-called *Garman-Klass estimator*, see [30], which is an unbiased estimator having minimum variance in the class of quadratic estimators. The Garman-Klass estimator is given by

$$\hat{\sigma}^2_{GK} = \frac{1}{n\Delta} \sum_{i=1}^{n} \Big\{ 0.511(H_{\Delta i} - L_{\Delta i})^2 - 0.019(X_{\Delta i}(H_{\Delta i} + L_{\Delta i})^2 - 2H_{\Delta i}L_{\Delta i}) - 0.383X^2_{\Delta i}) \Big\}. \quad (1.5)$$

The estimators presented above have in common that they are moment-type estimators based on second-order moments of the triplet $(H_\Delta, L_\Delta, X_\Delta)$. Alternatively, the range-based maximum likelihood estimator (ML-estimator) of σ can also be determined. This is due to the fact that the joint densities of (H, L) and (H, L, X) in the Brownian model with drift, described by (1.1), are explicitly known. Series expansions for these densities are described in the article of Dominé [20] or in the book of Revuz and Yor [57]. Magdon-Ismail and Atiya, see [49], analyzed the asymptotic behavior of the ML-estimator in a simulation study, as the number of observations n tends to infinity. They compared a numerically determined ML-estimator of σ^2 to the estimators (1.2)-(1.5). According to their study, the maximum-likelihood estimator seems to be more efficient than the moment estimators. However, the maximum likelihood estimator is hard to analyze theoretically because of the highly non-linear dependence of the joint density of (H, L, X) on the diffusion coefficient σ.

The aim of the present thesis is to derive statistical results similar to the ones above for diffusion models that are more general than Brownian motion. One encounters difficulties since the joint density $f(t, x, h, l, y)$ of (H_t, L_t, X_t), conditional on $X_0 = x$, usually has no explicit representation. It turns out that it can be determined by means of numerical methods for partial differential equations, but an approximation is hard to obtain. Therefore it is irrational to engage with the analysis of ML-estimators. Instead, we focus our attention on the search of reasonable alternative inference methods in the range-based context. Our means of first choice are so-called *martingale estimating functions*. These estimating functions are inspired by the maximum likelihood approach and result in moment-like estimators, which are sometimes also called *quasi-likelihood estimators*. For discretely sampled diffusion processes, various classes of estimating functions have already been studied extensively. For some references, see the discussion of martingale estimating functions in the upcoming Paragraph "Preview of presented results". Our mission is to generalize and to modify these results in such a way that they are applicable for range-based observations.

During our investigations several theoretical results about the joint distribution of the triplet (H, L, X) are derived. They are necessary to analyze the statistical behavior of (H, L, X) and can be considered as auxiliary results. Of course, this is not the only raison d'être of these analytical tools. The distributional properties of (H, L, X) are also very interesting themselves.

Preview of presented results

Let $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ be sufficiently smooth function. Moreover, let $(B_t,\, t \geq 0)$ denote the standard Brownian motion of $\mathbb{R}$. Throughout the present thesis, we are going to consider one-dimensional, time-homogeneous diffusions X that are defined by stochastic differential equations of the following type

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x,\ t \geq 0. \tag{1.6}$$

We denote the running maximum and the running minimum of X at time t with

$$H_t = \sup_{0 \leq s \leq t} X_s \quad \text{and} \quad L_t = \inf_{0 \leq s \leq t} X_s. \tag{1.7}$$

The object of our thesis is to exploit the information the triplet (H, L, X) carries in order to derive reasonable estimators of the diffusion coefficient σ. The contents of the different chapters and the precise issues we are going to address are outlined in the sequel.

In Chapter 2 we give a brief overview of how stochastic analysis and partial differential equation theory are intertwined. Some very elementary results are quoted. In particular we concentrate on diffusions killed at the boundary of an open interval $\mathcal{D} \subset \mathbb{R}$. It is not difficult to show that the distributional properties of these so-called *killed diffusions* can be used to determine the joint probability of triplet (H, L, X). Some very elaborate results for killed diffusions can be found in the book of Stroock [68] or in the book of Kallenberg [41]. We review some of the known results and we add some important details and alternative proofs.

At the beginning of Chapter 3, in Section 3.2, we present an existence result for the joint density of the triplet (H, L, X). Malliavin calculus is required to prove the statements. We give a brief introduction to this topic that is tailored to our needs. As an application of the existence result, we show how the joint density can be calculated – at least theoretically. The methods mostly rely on the results of Chapter 2 and can be found in Section 3.3. Beside the methodical results concerning the joint density of (H, L, X), we also present some facts about the Laplace transform of first hitting time densities of one-dimensional diffusions. The presented properties are based on the work of Darling and Siegert [18] and it turns out that in some cases the respective Laplace transform has a very simple representation.
There are some very important cases for which the density can be derived explicitly. In the Brownian model this is possible. But the joint density of (H_t, L_t, X_t) can also be calculated for the Ornstein-Uhlenbeck process. These and some other examples, that are closely related to our findings, are displayed in Section 3.4.

The approaches presented in Chapter 2 and Chapter 3 are mostly based on partial differential equation theory and, except for some simple cases, it is not possible to find an explicit representation for the joint density of (H, L, X). Nevertheless, the presented existence result enables us to establish a theory for martingale estimating functions based on the observations $(H_{i\Delta}, L_{i\Delta}, X_{i\Delta})$, $i = 1, \ldots, n$, in the parameterized one-dimensional diffusion model

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = x_0,\ t \geq 0. \tag{1.8}$$

Here, we assume that the sampling frequency $\Delta > 0$ is fixed, and we denote the suprema and the infima on the equidistant observation intervals $((i-1)\Delta, i\Delta]$, $i = 1, \ldots, n$, with $H_{i\Delta} = \sup_{(i-1)\Delta \leq s \leq i\Delta} X_s$ and $L_{i\Delta} = \inf_{(i-1)\Delta \leq s \leq i\Delta} X_s$. Martingale estimating functions, or estimating functions in general, are inspired by the maximum likelihood approach. They result in moment-like estimators for the parameter θ. Concretely, in the case of equidistant discrete observations $(X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta})$, the estimating equation

$$U_n(\theta) = \sum_{i=1}^{n} \partial_\theta \log p(\Delta, X_{(i-1)\Delta}, X_{i\Delta};\theta) = 0 \tag{1.9}$$

is replaced by another equation

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta, X_{(i-1)\Delta}, X_{i\Delta};\theta) = 0. \tag{1.10}$$

The function $G_n(\theta)$ has to be chosen in a sensible way. If $G_n(\theta)$ is a martingale with respect to $\mathcal{F}(X_\Delta, \ldots, X_{n\Delta})$, we call $G_n(\theta)$ a *martingale estimating function*. An estimator $\hat{\theta}_n$ that solves the equation (1.10) is sometimes called a *quasi-likelihood estimator* and the function g in (1.10) must belong to an adequate class of functions $\mathcal{G}_n$. The most canonical choice is the class that consists of functions of the following type

$$\begin{aligned} g(\Delta, x, y;\theta) &= a_1(\Delta, x;\theta)\Big\{y - \mathbb{E}_\theta[X_\Delta \,|\, X_0 = x]\Big\} \\ &\quad + \ a_2(\Delta, x;\theta)\Big\{\Big(y - \mathbb{E}_\theta[X_\Delta \,|\, X_0 = x]\Big)^2 - \mathrm{Var}_\theta[X_\Delta \,|\, X_0 = x]\Big\}. \end{aligned} \tag{1.11}$$

The essential issue is to determine the optimal weights a_1^* and a_2^* in such a way that the corresponding estimating function $G_n^*(\theta)$ is closest to the unknown score function $U_n(\theta)$ in the sense of Godambe and Heyde, see [32]. Consistency and asymptotic normality have been proved for quasi-likelihood estimators inferred from (optimal) martingale estimating functions based on discrete observations. See the work of Bibby and Sørensen, [11] and [12], Kessler and Sørensen [46], Kessler [44] and [45], and Pedersen [53]. An overview of the asymptotical methods, as the number of observations n tends to infinity, can be found in the work of Sørensen, see [66]. A more recent paper that summarizes the existing results was written by Bibby, Jacobsen and Sørensen [9].
We give a very brief introduction to the general ideas of martingale estimating functions in the simple case of equidistant observations at the beginning of Chapter 4. Then we generalize the existing results by replacing the single equidistant observations $X_{i\Delta}$ by the

vectors $(H_{i\Delta}, L_{i\Delta}, X_{i\Delta})$, $i = 1, \dots, n$, and by replacing the transition density $p(\Delta, x, y; \theta)$ by the joint density $f(\Delta, x, h, l, y; \theta)$ of $(H_\Delta, L_\Delta, X_\Delta)$, conditional on $X_0 = x$. An interesting example for a generalized martingale estimating function is given by the following quadratic martingale estimating function

$$g_{qua,gen}(\Delta, x, h, y; \theta) = \sum_{j=1}^{3} a_j(\Delta, x; \theta) k_j(\Delta, x, h, y; \theta) \tag{1.12}$$

with

$$\begin{aligned} k_1(\Delta, x, h, y; \theta) &= \Big[h - \mathbb{E}_\theta[H_\Delta \,|\, X_0 = x]\Big]^2 - \mathrm{Var}_\theta \Big[H_\Delta \,|\, X_0 = x\Big], \\ k_2(\Delta, x, h, y; \theta) &= \Big[y - \mathbb{E}_\theta[X_\Delta \,|\, X_0 = x]\Big]^2 - \mathrm{Var}_\theta \Big[X_\Delta \,|\, X_0 = x\Big], \\ k_3(\Delta, x, h, y; \theta) &= \Big[h - \mathbb{E}_\theta[H_\Delta \,|\, X_0 = x]\Big]\Big[y - \mathbb{E}_\theta[X_\Delta \,|\, X_0 = x]\Big] - \mathrm{Cov}_\theta \Big[H_\Delta,\, X_\Delta \,|\, X_0 = x\Big]. \end{aligned} \tag{1.13}$$

This function appears several times during our discussions. It is only a very special case. Many other estimating functions are imaginable. For example, any properly normalized polynomial in the expressions

$$\Big[h - \mathbb{E}_\theta[H_\Delta \,|\, X_0 = x]\Big], \quad \Big[l - \mathbb{E}_\theta[L_\Delta \,|\, X_0 = x]\Big] \quad \text{and} \quad \Big[y - \mathbb{E}_\theta[X_\Delta \,|\, X_0 = x]\Big] \tag{1.14}$$

can be used to construct estimating functions.

On several technical assumptions, we prove consistency and we analyze the asymptotical properties of estimators inferred from generalized estimating functions as the number of observations n tends to infinity, see Section 4.2.1. Particularly, we are able to state a result about asymptotic normality. Moreover, in Section 4.2.3, we make use of Godambe and Heyde's optimality theory to establish optimality criteria for our generalized model. As for the ordinary case, this leads to optimal weights $a_i^*(\Delta, x; \theta)$ for the generalized estimating functions. For several classes of estimating functions, these optimal weights are calculated explicitly in Section 4.3.

The results of Chapter 4 are highly theoretical though, since the joint probability or the joint density of (H_t, L_t, X_t), for a diffusion process X starting in $x \in \mathbb{R}$, cannot be calculated explicitly in general. This is the reason why we have to think about how our statistical model can be simplified in a reasonable way. The most obvious possibility is to approximate the optimal estimating functions. However, this is not straightforward. Some auxiliary results are necessary. As a first step, in Chapter 5, we try to find an expansion of $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$. Here, $g : \mathbb{R}^3 \to \mathbb{R}$ denotes a sufficiently smooth function that does not grow too fast. Our approach makes use of elementary estimates and, as a result, we obtain an expansion of $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$ including powers of 2. Furthermore, this expansion can only be calculated for diffusion processes X with constant diffusion coefficient $\sigma > 0$. Note that this is not a real restriction, since the stochastic differential equation (1.6) can be transformed into a model where the diffusion coefficient equals 1. This modification is done by the so-called *Lamperti transform*. The transformed model is basically equivalent to the original model. More information can be found in Chapter 5.

For the case $\sigma \equiv 1$, our second order expansion with respect to $\sqrt{t}$ is given by

$$\begin{aligned}
\mathbb{E}_x[g(H_t, L_t, X_t)] = g(x,x,x) & \\
& + g_{1,0,0}(x,x,x)\sqrt{\frac{2}{\pi}}\sqrt{t} + g_{1,0,0}(x,x,x)\frac{1}{2}\mu(x)t + g_{2,0,0}(x,x,x)\frac{1}{2}t \\
& - \ g_{0,1,0}(x,x,x)\sqrt{\frac{2}{\pi}}\sqrt{t} + g_{0,1,0}(x,x,x)\frac{1}{2}\mu(x)t + g_{0,2,0}(x,x,x)\frac{1}{2}t \\
& + \ (1 - 2\log 2)g_{1,1,0}(x,x,x)t \\
& + \ \frac{1}{2}g_{1,0,1}(x,x,x)t + \frac{1}{2}g_{0,1,1}(x,x,x)t \\
& + \ g_{0,0,1}(x,x,x)\mu(x)t + \frac{1}{2}g_{0,0,2}(x,x,x)t \\
& + \ O(t^{3/2}).
\end{aligned} \tag{1.15}$$

In the above equation we use the notation $g_k(x,x,x)$ to denote the partial derivatives

$$g_k(x,x,x) = \frac{\partial^{|k|}}{\partial h^{k_1}\partial l^{k_2}\partial y^{k_3}} g(h,l,y)\Big|_{(h,l,y)=(x,x,x)}, \quad \text{with } k = (k_1,k_2,k_3) \in \mathbb{N}_0^3. \tag{1.16}$$

And, of course, one immediately discovers the associated moments of

$$\left(\sup_{0\le s\le 1} B_s, \inf_{0\le s\le 1} B_s, B_1\right). \tag{1.17}$$

It is well known that, for example,

$$\mathbb{E}\left[\sup_{0\le s\le 1} B_s\right] = \sqrt{\frac{2}{\pi}}, \quad \mathbb{E}\left[\left(\sup_{0\le s\le 1} B_s\right)^2\right] = 1, \quad \text{or} \quad \mathbb{E}\left[\left(\sup_{0\le s\le 1} B_s\right)\cdot B_1\right] = \frac{1}{2}. \tag{1.18}$$

We also know that

$$\mathbb{E}\left[\left(\sup_{0\le s\le 1} B_s\right)\left(\inf_{0\le s\le 1} B_s\right)\right] = 1 - 2\log 2. \tag{1.19}$$

The latter formula was proved by Rogers and Shepp [62]. The terms

$$g_{1,0,0}(x,x,x)\frac{1}{2}\mu(x)t \quad \text{and} \quad g_{0,1,0}(x,x,x)\frac{1}{2}\mu(x)t \tag{1.20}$$

strike the eye. They are more difficult to explain. For more details see Chapter 5, especially see Theorem 5.2.3.5, where the overall result is stated.

The presented expansion is already sufficient to state different results about *small-Δ-optimal martingale estimating functions* in our generalized model. *Small-Δ-optimality* is a notion that has been established by Jacobsen, see [37] and [38]. Also see Bibby, Jacobsen and Sørensen [9] for a summary of Jacobsen's results. Jacobsen takes an expansion of the transition operator $\mathbb{E}_x[g(X_\Delta)]$ into account to find approximately optimal martingale estimating functions on small observation intervals $((i-1)\Delta, i\Delta]$, $i = 1, \ldots, n$. Thereby the label small-Δ-optimality is explained and, in the ordinary model, this concept amounts to replacing the optimal weights a_1^* and a_2^* in formula (1.11) by their first order approximations. For a fixed sample size n, a lower bound for the variance of the estimating functions can be found that is uniform as $\Delta \to 0$. Besides, the approximations of the weights a_1^* and a_2^* can be chosen in such a way that the aforementioned asymptotic lower bound of the variance is attained. Here our above expansion of $\mathbb{E}_x[g(H_\Delta, L_\Delta, X_\Delta)]$ comes into play. It allows for a generalization of the concept of small-Δ-optimality. An introduction to small-Δ-optimal martingale estimating functions for equidistant observations $X_{i\Delta}$, $i = 1, \ldots, n$, and our generalized results that are based on the observations $(H_{i\Delta}, L_{i\Delta}, X_{i\Delta})$, $i = 1, \ldots, n$, can be found in Chapter 6. Different classes of linear and quadratic generalized martingale estimating functions are discussed in Section 6.3.3. For these classes small-Δ-optimality results are derived and the resulting estimators are assessed by a comparison with the corresponding estimators stemming from the ordinary small-Δ-optimal model that relies on equidistant observations alone. Particularly, we discover that we can benefit from our generalized concept when it comes to estimating a parameter θ in the diffusion coefficient $\sigma(\cdot;\theta)$. These findings contrast with the case of drift estimation, where incorporating the values $(H_{i\Delta}, L_{i\Delta})$ does not decrease the small-Δ-asymptotic lower bound for the variance of the estimators of $\mu(\cdot,\theta)$.

As an application of the theoretical results we present a case study for an Ornstein-Uhlenbeck process at the end of Chapter 6. The simulations support our theoretical findings. If we want to estimate a parameter θ in the diffusion coefficient $\sigma(\cdot,\,\theta)$, there is a gain in efficiency originating from a replacement of the ordinary martingale estimating functions with equidistant observations $X_{i\Delta}$, by generalized martingale estimating functions, that are based on the knowledge of the values $(H_{i\Delta}, L_{i\Delta}, X_{i\Delta})$. But if the observation intervals $((i-1)\Delta, i\Delta]$ are relatively large, the small-Δ-optimal estimators of $\sigma(\cdot,\theta)$, which are constructed by means of an expansion of $\mathbb{E}_x[g(H_\Delta, L_\Delta, X_\Delta)]$, do not yield reliable results. This is due to the fact that we work with expansions of low order yielding poor approximations to the moments for large values of Δ. Therefore it is highly desirable to know an overall expansion – or at least a higher order expansion – of $\mathbb{E}_x[g(H_\Delta, L_\Delta, X_\Delta)]$ with respect to $\sqrt{\Delta}$, so that more accurate estimators can be determined.

In Chapter 7 we tackle the problem of determining an expansion of $\mathbb{E}_x[g(H_t, X_t)]$ with respect to $\sqrt{t}$. We are able to obtain an overall expansion for a certain class of diffusion processes and these results carry over to $\mathbb{E}_x[g(L_t, X_t)]$, where the running maximum H is replaced by the running minimum L. The problem of expanding $\mathbb{E}_x[g(H_t, L_t, X_t)]$ is completely neglected. But our research is a first important step on the way to determining the exact behavior of the triplet (H_t, L_t, X_t).

We will achieve the goal of expanding $\mathbb{E}_x[g(H_t, X_t)]$ via a detour. Initially, we concentrate on the study of first hitting times of pinned diffusions. This means we want to find an expansion of the expression

$$\mathbb{P}_x[\tau_h \leq t \,|\, X_t = y], \tag{1.21}$$

with respect to t, where $\tau_h = \inf\{t > 0 \,|\, X_t = h\}$ and $x, y \leq h$. The asymptotics of hitting times for pinned processes as $t \to 0$ have recently been studied by several authors. Important contributions are the work of Peskir [55] or the work of Borovkov and Downes [17]. For our purposes the research of Baldi and Caramellino [6] turns out to be most relevant. Baldi and Caramellino investigate the asymptotical properties of the quantity (1.21) by means of large deviation techniques. Their results can be considered as a first order expansion of (1.21) with respect to t. But large deviation principles do not allow for a generalization of their result. A completely new approach is necessary to find an overall expansion of (1.21). Some rather elementary transformations show that the following functional of the Brownian bridge

$$\mathbb{E}_x\left[\exp\left(-\frac{t}{2}\int_0^1 \beta\Big(\sqrt{t}B_u\Big)du\right)\mathbb{1}_{\{\sup_{0\leq u\leq 1}\sqrt{t}B_u \geq h\}} \,\Big|\, B_1 = y\right] \tag{1.22}$$

is the key component of the expression (1.21). In formula (1.22) B denotes the standard one-dimensional Brownian motion and $\beta : \mathbb{R} \to \mathbb{R}$ is a function that depends on the coefficients μ and σ of the diffusion X. We derive an overall expansion of (1.22) which has the following form

$$\exp\left(-2\frac{(h-x)(h-y)}{t}\right)\left\{1 + \sum_{i=1}^{\infty} t^i \tilde{\phi}_i(1, x, h, y)\right\}. \tag{1.23}$$

The coefficients $\tilde{\phi}_i$ in the expansion of (1.23) are described by a system of partial differential equations that are recursively defined and that must clearly depend on the function β. This approach is inspired by the methods of Fleming and James [25], who combine different PDE techniques in order to find an expansion of the rather general Feynman-Kac formula

$$\mathbb{E}_x\left[\exp\left(\int_0^t g_1(X_s)ds\right)\right] + \mathbb{E}_x\left[\int_0^t \exp\left(\int_0^s g_1(X_u)du\right) g_2(X_s)ds\right], \tag{1.24}$$

if X is a general diffusion processes and for sufficiently smooth functions $g_1, g_2 : \mathbb{R} \to \mathbb{R}$. Once the expansion of (1.22) is determined, it is relatively simple to retransform the series into an expansion of the first hitting time probability of X pinned at $X_t = y$.
The expansion of (1.22) can be found in Chapter 7.4.2, see especially Corollary 7.4.1.3.
The retransformation of the expansion of (1.22) into an expansion of $\mathbb{P}_x[\tau_h \leq t \,|\, X_t = y]$

is conducted in Section 7.5. The overall result is stated in Theorem 7.5.0.4.

By the result of Chapter 7 it is possible to derive an overall expansion of the joint distribution of (H_t, X_t). This follows directly from the fact that, for a measurable set $A \in \mathcal{B}(\mathbb{R})$, the following relation holds:

$$\mathbb{P}_x[H_t \geq h,\ X_t \in A] = \int_A \mathbb{P}_x\left[\tau_h \leq t \,|\, X_t = y\right] \mathbb{P}_x[X_t \in dy]. \tag{1.25}$$

In order to give a concrete example, in Chapter 8, we explicitly calculate a fourth order expansion of $\mathbb{E}_x[g(H_t, X_t)]$, with respect to $\sqrt{t}$, for a diffusion model with diffusion coefficient $\sigma \equiv 1$. We benefit from the fact that several expansions for the transition density $p(t, x, y)$ of X have already been derived. The expansion of $p(t, x, y)$ we are going to use was calculated by Aït-Sahalia, see [3] or [4]. Our ultimate result, the fourth order expansion, is given by

$$\begin{aligned}
&\mathbb{E}_x[g(H_t, X_t)] \\
&= g(x,x) \\
&+ g_{(1,0)}(x,x)\left\{\sqrt{\frac{2}{\pi}}\sqrt{t} + \frac{1}{2}\mu(x)t + \frac{1}{4}t^2\left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x)\right) + \frac{1}{3}\frac{1}{\sqrt{2\pi}}t^{3/2}\left(\mu'(x) + \mu(x)^2\right)\right\} \\
&+ g_{(0,1)}(x,x)\left\{\mu(x)t + \frac{1}{2}t^2\left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x)\right)\right\} \\
&+ g_{(2,0)}(x,x)\frac{1}{2}\left\{t + \frac{4}{3}\sqrt{\frac{2}{\pi}}\mu(x)t^{3/2} + \frac{1}{2}t^2(\mu'(x) + \mu(x)^2)\right\} \\
&+ g_{(0,2)}(x,x)\frac{1}{2}\left\{t + t^2(\mu'(x) + \mu(x)^2)\right\} \\
&+ g_{(1,1)}(x,x)\left\{\frac{1}{2}t + \frac{4}{3}\sqrt{\frac{2}{\pi}}\mu(x)t^{3/2} + \frac{1}{2}t^2(\mu'(x) + \mu(x)^2)\right\} \\
&+ g_{(3,0)}(x,x)\frac{1}{6}\left\{2\sqrt{\frac{2}{\pi}}t^{3/2} + \frac{9}{4}\mu(x)t^2\right\} + g_{(0,3)}(x,x)\frac{1}{2}\mu(x)t^2 \\
&+ g_{(2,1)}(x,x)\frac{1}{2}\left\{\frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2} + 2\mu(x)t^2\right\} + g_{(1,2)}(x,x)\frac{1}{2}\left\{\frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2} + \frac{3}{2}\mu(x)t^2\right\} \\
&+ g_{(4,0)}(x,x)\frac{1}{8}t^2 + \frac{1}{8}g_{(0,4)}(x,x)t^2 + g_{(3,1)}(x,x)\frac{3}{8}t^2 + g_{(1,3)}(x,x)\frac{1}{4}t^2 + g_{(2,2)}(x,x)\frac{1}{2}t^2 \\
&+ O(t^{5/2}).
\end{aligned} \tag{1.26}$$

A comparison of (1.26) with (1.15) shows that the coefficients belonging to $\sqrt{t}$ and to t are the same in both expansions. The higher order terms are harder to interpret. For example, the expectations

$$\mathbb{E}\left[\left(\sup_{0\leq s\leq 1} B_s\right)^4\right] = 3 \quad \text{and} \quad \mathbb{E}\left[\left(\sup_{0\leq s\leq 1} B_s\right)^2 \cdot B_1^2\right] = 2, \tag{1.27}$$

are directly related to the terms

$$g_{(4,0)}(x,x)\frac{1}{8}t^2 \quad \text{and} \quad g_{(2,2)}(x,x)\frac{1}{2}t^2. \tag{1.28}$$

Note that $4! = 24$ and $2! \cdot 2! = 4$. But, apart from such simple examples, the expressions in (1.26) are not all completely obvious. Besides, it is also noteworthy that the above expansion is not an expansion in the classical sense. The highest order term in the Taylor expansion of a smooth function g only depends on the highest order derivative of g. Here, this is apparently not the case. For more information about our expansion we make reference to Theorem 8.3.3.1. Finally let us note that the expectation $\mathbb{E}_x[g(H_t, X_t)]$ can be expanded further. But the calculations are tedious since every coefficient has to be calculated individually. Nevertheless, this result constitutes a large improvement compared to the elementary approach of Chapter 5.

At the beginning of each chapter, we will summarize technical results that are relevant to our analyses. Additionally, there will be an introduction to the respective topics with more references to related results in literature. The relevant notations will be explained when they appear for the first time within a chapter. Furthermore, each chapter is written in such a way that it can be read separately.

2 General Concepts for Diffusion Models

2.1 Introduction & Motivation

A diffusion process – or simply diffusion – on an interval $\mathcal{D} \subseteq \mathbb{R}$ is a Markov process with continuous paths and state space $\mathcal{D}$ having the strong Markov property. Let $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ denote the space of continuous functions on $\mathbb{R}_+$ taking values in $\mathbb{R}$, and let $X = (X_t,\, t \geq 0)$ denote the coordinate process on Ω defined by $X_t(\omega) = \omega(t)$ for $\omega \in \Omega$. We assume that the space Ω is endowed with the σ-algebra $\mathcal{F} = \sigma\{X_t\,;\, 0 \leq t < \infty\}$. Finally, we denote by $\mathbb{P}_x$ the Markov measure on $(\Omega, \mathcal{F})$ making $\mathbb{P}_x[X_0 = x] = 1$. The corresponding expectation operator is denoted by $\mathbb{E}_x$.

The processes we are especially interested in are diffusions defined by the following time-homogeneous stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x \in \mathbb{R},\ t \geq 0, \tag{2.1}$$

where the driving process $(B_t,\, t \geq 0)$ is a standard Brownian motion of $\mathbb{R}$. Note that a sufficient criterion for the existence of a weak solution $X = (X_t,\, t \geq 0)$ of (2.1) is that $\mu : \mathbb{R} \to \mathbb{R}$ is measurable and bounded, and that $\sigma : \mathbb{R} \to \mathbb{R}_+$ is continuous. See, for example, the book of Stroock and Varadhan [69].

For $t > 0$, we want to consider the triplet (H_t, L_t, X_t), where

$$H_t = \sup_{0 \leq s \leq t} X_s \quad \text{and} \quad L_t = \inf_{0 \leq s \leq t} X_s. \tag{2.2}$$

In Section 3.2 we will derive an existence result for the joint density of (H_t, L_t, X_t). In order to prove this result, we will have to impose several regularity conditions. Concretely, we will assume that the coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ are continuously differentiable functions, with uniformly bounded first derivatives, and that σ is uniformly elliptic. In this case μ and σ satisfy a global Lipschitz condition, i.e. there is some constant $K > 0$ such that

$$\big|\sigma(x) - \sigma(y)\big| + \big|\mu(x) - \mu(y)\big| \leq K|x - y|, \quad \forall x, y \in \mathbb{R}. \tag{2.3}$$

It is a well known result that, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a Brownian motion B exists, the above Lipschitz condition even implies the existence of a unique solution to (2.1) in the strong sense. See for example Stroock and Varadhan [69].

Yet, the aim of the present chapter is to establish the foundations for an analysis of one-dimensional diffusion processes X that are killed at the boundary of an interval. As we will see in Chapter 3, the transition density of this very special kind of diffusions can be used to calculate the joint density of (H_t, L_t, X_t). The behavior of killed diffusions has been extensively analyzed. The main references for probabilists are the book of Stroock, see [68], or the books of Friedman, see [28] and [29]. Different existing results are quoted and adjusted to our purposes. Moreover, for some of the results, we give alternative proofs that are very intuitive.

The structure of the present chapter is as follows. At the beginning we present the martingale problem and some facts about semigroup theory in order to introduce the relevant notations and concepts. The proofs for these results can be found in the book of Kallenberg [41], for example. In Section 2.3, we show how the transition density of a diffusion killed at the boundary of an interval is related to Kolmogorov's equation with Dirichlet boundary conditions.

2.2 The martingale problem and some semigroup theory

In the present section, we briefly present some well known results. The following concepts and the missing proofs can all be found in the book of Kallenberg [41]. Let $m \in \mathbb{N}$. For an open set $S \subseteq \mathbb{R}^m$ and for a multi-index $k = (k_1, \ldots, k_m) \in \mathbb{N}_0^m$, we denote the space of functions $g : S \to \mathbb{R}$, for which the partial derivatives

$$\frac{\partial^{|k|}}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}} g(x_1, \ldots, x_m) \tag{2.4}$$

exist and are continuous, with $\mathcal{C}^k(S, \mathbb{R})$. Here, $|k| = k_1 + \ldots + k_m$. Moreover, $\mathcal{C}_c^k(S, \mathbb{R})$ denotes the space of functions belonging to $\mathcal{C}^k(S, \mathbb{R})$ and having compact support on S, and $\mathcal{C}_b^k(S, \mathbb{R})$ denotes the space of functions belonging to $\mathcal{C}^k(S, \mathbb{R})$ and being bounded on S. We will usually write $\mathcal{C}(S, \mathbb{R})$, $\mathcal{C}_c(S, \mathbb{R})$ and $\mathcal{C}_b(S, \mathbb{R})$ instead of $\mathcal{C}^0(S, \mathbb{R})$, $\mathcal{C}_c^0(S, \mathbb{R})$ and $\mathcal{C}_b^0(S, \mathbb{R})$.

Definition 2.2.0.1. Let $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ be measurable functions and for a function $f : \mathbb{R} \to \mathbb{R}$ that is twice continuously differentiable let $\mathcal{A}$ denote the operator

$$\mathcal{A}f(x) := \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}f(x) + \mu(x)\frac{\partial}{\partial x}f(x). \tag{2.5}$$

A probability measure $\mathbb{P}$ on the path space $(\Omega, \mathcal{F})$ is called a solution to the local martingale problem for (μ, σ) if

$$M^f(t) := f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s)ds, \quad t \geq 0, \tag{2.6}$$

is a local martingale under $\mathbb{P}$ for all functions $f \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$.

Remark 2.2.0.2. If μ and σ are bounded, $\mathbb{P}$ even solves the martingale problem for which M^f is required to be a proper martingale for every $f \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$.

Theorem 2.2.0.3. *The stochastic differential equation* (2.1) *has a weak solution if and only if a solution to the local martingale problem* (μ, σ) *exists. In this case the law* $\mathbb{P}_x$ *of* X *on the path space equals the solution of the local martingale problem.*

Corollary 2.2.0.4. *A stochastic differential equation has a (in distribution) unique weak solution if and only if the corresponding local martingale problem is uniquely solvable, given some initial distribution.*

Let us turn our attention to semigroup theory. First, we state the following lemma.

Lemma 2.2.0.5. *A family* $(k_t)_{t \geq 0}$ *of probability kernels from* $\mathbb{R}$ *to* $\mathcal{B}(\mathbb{R})$ *satisfies the Chapman-Kolmogorov relation* $k_{s+t} = k_s\, k_t$ *for all* $t, s \geq 0$*, if and only if the transition operators*

$$T_t f(x) = \int f(y) k_t(x, dy), \;\; f : \mathbb{R} \to \mathbb{R} \;\; \text{bounded and measurable,} \tag{2.7}$$

form a semigroup, that is $T_t \circ T_s = T_{t+s}$ *holds for all* $t, s \geq 0$.

Definition 2.2.0.6. If the operators $(T_t)_{t \geq 0}$ satisfy

(a) $T_t f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$ for all $f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$, and

(b) $\lim_{h \to 0} T_h f(x) = f(x)$ for all $f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$, $x \in \mathbb{R}$,

then $(T_t)_{t \geq t}$ is called a *Feller semigroup*.

Theorem 2.2.0.7. *A Feller semigroup* $(T_t)_{t \geq t}$ *is uniquely determined by its infinitesimal generator* $\mathcal{A} : \mathrm{dom}(\mathcal{A}) \subset \mathcal{C}(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$ *with*

$$\mathcal{A} f := \lim_{h \to 0} \frac{T_h f - f}{h}, \tag{2.8}$$

and

$$\mathrm{dom}(\mathcal{A}) = \left\{ f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \Big| \lim_{h \to 0} \frac{T_h f - f}{h} \; \textit{exists} \right\}. \tag{2.9}$$

Moreover, the semigroup uniquely defines the probability kernels and thus it determines the distribution of the associated Markov process (which is called Feller process).

Corollary 2.2.0.8. *We have for* $f \in \mathrm{dom}(\mathcal{A})$,

$$\frac{d}{dt} T_t f = \mathcal{A} T_t f = T_t \mathcal{A} f. \tag{2.10}$$

Theorem 2.2.0.9 (Hille-Yosida)**.** *Let* $\mathcal{A}$ *be a closed linear operator on* $\mathcal{C}(\mathbb{R}, \mathbb{R})$ *with dense domain* $\mathrm{dom}(\mathcal{A})$. *Then* $\mathcal{A}$ *is the generator of a Feller semigroup if and only if*

1. *the range of* $\lambda_0(\mathrm{Id} - \mathcal{A})$ *is dense in* $\mathcal{C}(\mathbb{R}, \mathbb{R})$ *for some* $\lambda_0 > 0$*;*
2. *if for some* $x \in \mathbb{R}$ *and* $f \in \mathrm{dom}(\mathcal{A})$*,* $f(x) \geq 0$ *and* $f(x) = \max_{y \in \mathbb{R}} f(y)$ *then* $\mathcal{A}f(x) \leq 0$ *follows (positive Maximum principle).*

Hille-Yosida's theorem will not be used in the sequel. We mentioned it for the sake of completeness only. The next theorem states an important property of the infinitesimal generator for stochastic differential equations.

Theorem 2.2.0.10. *Let* μ *and* σ *be measurable coefficients such that the local martingale problem for* (μ, σ) *has a unique solution* $\mathbb{P}_x$ *for all initial distributions* δ_x*,* $x \in \mathbb{R}$*, then the Markov measures* $(\mathbb{P}_x)_{x \in \mathbb{R}}$ *solving the martingale problem for* (μ, σ) *give rise to a Feller semigroup* $(T_t)_{t \geq 0}$*. Any function* $f \in \mathcal{C}_c^2(\mathbb{R}, \mathbb{R})$ *lies in* $\mathrm{dom}(\mathcal{A})$ *and fulfills*

$$\mathcal{A}f(x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}f(x) + \mu(x)\frac{\partial}{\partial x}f(x). \tag{2.11}$$

Theorem 2.2.0.11 (Dynkin's formula)**.** *Assume that* μ *and* σ *are measurable, locally bounded and such that the stochastic differential equation* (2.1) *with time homogeneous coefficients has a (in distribution) unique weak solution. Then for all* $x \in \mathbb{R}$*,* $f \in \mathcal{C}_c^2(\mathbb{R}, \mathbb{R})$ *and all bounded stopping times* τ *we have*

$$\mathbb{E}_x\left[f(X_\tau)\right] = f(x) + \mathbb{E}_x\left[\int_0^\tau \mathcal{A}f(X_s)ds\right]. \tag{2.12}$$

2.3 Killed diffusions

Let us consider a diffusion process X that satisfies the stochastic differential equation (2.1). Throughout the present section, we assume that the following minimal assumption is satisfied - if not otherwise specified.

Condition 2.3.0.12. *We assume that* μ *and* σ *belong to* $\mathcal{C}_b^1(\mathbb{R}, \mathbb{R})$ *with uniformly bounded first derivatives, and that* σ *is uniformly elliptic.*

Remark 2.3.0.13. The above assumption is relatively strong. It is necessary to prove some of the regularity results of this section. In later chapters, we will be able to weaken the boundedness condition for μ and σ. From Chapter 3, we will mostly work on the condition of linear growth.

Let $\mathcal{D} \subset \mathbb{R}$ be an open interval and let us introduce the first exit time $\zeta_{\mathcal{D}}$ of the diffusion process X which is defined by

$$\zeta_{\mathcal{D}} = \inf\{t > 0 \mid X_t \in \mathcal{D}^c\}. \tag{2.13}$$

From this stopping time a new process can be inferred, namely the process that coincides with X on the interior of $\mathcal{D}$ and that is killed if X hits the boundary $\partial\mathcal{D}$. We denote this process by $X^{\mathcal{D}}$. Formally it is defined by $X_t^{\mathcal{D}} = X_t$ if $t < \zeta_{\mathcal{D}}$ and $X_t^{\mathcal{D}} = \partial$ if $t \geq \zeta_{\mathcal{D}}$, where ∂ is an isolated point that is also called the cemetery state. Thus the state space of the process $X^{\mathcal{D}}$ is $\mathcal{D} \cup \{\partial\}$ and the associated semigroup of operators $(T_t^{\mathcal{D}},\, t \geq 0)$ is defined by

$$T_t^{\mathcal{D}} f(x) = \mathbb{E}_x[f(X_t),\, t < \zeta_{\mathcal{D}}] = \mathbb{E}_x\left[f(X_t)\mathbb{1}_{\{t<\zeta_{\mathcal{D}}\}}\right], \tag{2.14}$$

for $\mathcal{B}(\mathbb{R})$-measurable and bounded functions $f : \mathbb{R} \to \mathbb{R}$. This semigroup is sometimes called the *killed semigroup*.

The aim of this paragraph is to analyze the properties of the transition probability density $p^{\mathcal{D}}$ of the process $X^{\mathcal{D}}$. By definition $p^{\mathcal{D}}$ satisfies

$$\mathbb{P}_x[X_t^{\mathcal{D}} \in B] = \int_B p^{\mathcal{D}}(t, x, y) dy, \tag{2.15}$$

for all $x \in \mathcal{D}$ and for any $B \in \mathcal{B}(\mathcal{D})$. But first, we have to find conditions on which such a function exists. Therefore, let us note that for any measurable function $f \geq 0$ and for all $x \in \mathcal{D}$ we have $T_t^{\mathcal{D}} f(x) \leq T_t f(x) = \mathbb{E}_x[f(X_t)]$. Thus, if the process X has a transition density $p(t, x, y)$ with respect to the Lebesgue measure, it follows by the Radon-Nykodim theorem that the killed process $X^{\mathcal{D}}$ also has a transition density which satisfies (2.15). Additionally, it is obvious that the following estimate

$$0 \leq p^{\mathcal{D}}(t, x, y) \leq \mathbb{1}_{\mathcal{D}}(y) p(t, x, y) dy \tag{2.16}$$

holds. A sufficient criterion, that ensures the existence of a transition density p of X, is that μ and σ are globally Lipschitz functions of at most linear growth – thus Condition 2.3.0.12 implies the existence of a density. The proof usually either relies completely on analytical results or on Malliavin calculus. For a concise proof see e.g. Nualart [50].

By the strong Markov property one obtains that for any $t > 0$, any Borel set $B \in \mathcal{B}(\mathcal{D})$ and any $x \in \mathcal{D}$,

$$\mathbb{P}_x[X_t \in B] = \mathbb{P}_x[X_t \in B, t \leq \zeta_{\mathcal{D}}] + \mathbb{E}_x[T_{t-\zeta_{\mathcal{D}}} \mathbb{1}_B(X_{\zeta_{\mathcal{D}}}),\, t > \zeta_{\mathcal{D}}], \tag{2.17}$$

and consequently $p^{\mathcal{D}}$ satisfies

$$p^{\mathcal{D}}(t, x, y) = p(t, x, y) - \mathbb{E}_x[p(t - \zeta_D, X_{\zeta_{\mathcal{D}}}, y)\, \mathbb{1}_{\{t>\zeta_{\mathcal{D}}\}}], \tag{2.18}$$

for all $x, y \in \mathcal{D}$ and for all $t \geq 0$. This particular form of the transition density suggests that the function $x \mapsto p^{\mathcal{D}}(t, x, y)$ must vanish at the boundary of $\mathcal{D}$. A concise proof for this fact will be given in the sequel. Moreover, we will state the continuity of $p^{\mathcal{D}}$. A proof for the Brownian case can be found in the book of Kallenberg [41], see Theorem 24.7 therein, or in the book of Karatzas and Shreve [43].

First, let us state an auxiliary lemma.

Lemma 2.3.0.14. *Let $\mathcal{D} \subset \mathbb{R}$ denote a bounded open interval and let X be a diffusion* *that is defined on* $\mathbb{R}$ *and satisfies the stochastic differential equation* (2.1)*. We assume that the coefficients* μ *and* σ *satisfy Condition 2.3.0.12. Then, for* $b \in \partial\mathcal{D}$*, we have*

$$\mathbb{P}_b[\zeta_{\mathcal{D}} = 0] = 1, \tag{2.19}$$

where $\zeta_{\mathcal{D}}$ *is defined by* (2.13)*.*

Proof. On the Condition 2.3.0.12, the process X is a regular diffusion process. Note that a one-dimensional diffusion is called regular if any point can be hit from any other point in finite time. Let $b \in \partial\mathcal{D}$. In this case, the stopping time $T_b = \inf\{t > 0 | X_t > b\}$ satisfies $\mathbb{P}_b[T_b = 0] = 1$. This was proved by Rogers and Williams, see [60], Chapter V.46. By the fact that $0 \leq \zeta_D \leq T_b$, it follows that $\mathbb{P}_b[\zeta_{\mathcal{D}} = 0] = 1$ as well, which is the desired result. □

The next result is crucial. It shows that, for a bounded open interval $\mathcal{D}$, the function $x \mapsto p^{\mathcal{D}}(t, x, y)$ vanishes at the boundary $\partial\mathcal{D}$.

Theorem 2.3.0.15. *Suppose that the coefficients* μ *and* σ *satisfy Condition 2.3.0.12. For any bounded open interval* $\mathcal{D} \in \mathbb{R}$*, the function* $(t, x, y) \mapsto p^{\mathcal{D}}(t, x, y)$ *is continuous on* $(0, \infty) \times \mathcal{D}^2$*. If* $b \in \partial\mathcal{D}$*, then* $p^{\mathcal{D}}(t, x, y) \to 0$ *as* $x \to b$ *for fixed* $t > 0$ *and* $y \in \mathcal{D}$*.*

Proof. As we have already mentioned, for Brownian motion this result can be found in the book of Kallenberg [41], Theorem 24.7. Here some additional considerations are necessary.

The proof of Theorem 5.2.8 in the book of Stroock [68] shows that, on the assumptions made about the differentiability of the coefficients μ and σ, the function $(t, x, y) \mapsto p^{\mathcal{D}}(t, x, y)$ is continuous on $(0, \infty) \times \mathcal{D}^2$. Besides, Friedman [28] proved that, for any $T > 0$, the function $(t, x, y) \mapsto p(t, x, y)$ is continuous on $(0, T] \times \mathbb{R}^2$, where p denotes the ordinary transition density of X, and that there exist positive constants λ_1 and λ_2, depending on μ and σ only, such that

$$p(t, x, y) \leq \frac{\lambda_1}{\sqrt{t}} \exp\left(-\frac{|y - x|^2}{\lambda_2 t}\right), \quad \forall (t, x, y) \in [0, T] \times \mathbb{R}^2. \tag{2.20}$$

Let $t > 0$ and $y \in \mathcal{D}$ be fixed. By Lemma 2.3.0.14, for $b \in \partial\mathcal{D}$, we have $\mathbb{P}_b[\zeta_{\mathcal{D}} = 0] = 1$. Let $x \in \mathcal{D}$, then by the Markov property

$$\mathbb{P}_x[\zeta_{\mathcal{D}} > t] \leq \mathbb{P}_x[\zeta_{\mathcal{D}} \circ \theta_s > t - s] = \mathbb{E}_x[\mathbb{P}_{X_s}[\zeta_{\mathcal{D}} > t - s]]. \tag{2.21}$$

The right hand side is continuous in x, which follows by the continuity of the transition density of X and by dominated convergence. Thus we find

$$\limsup_{x \to b} \mathbb{P}_x[\zeta_{\mathcal{D}} > t] \leq \mathbb{E}_b[\mathbb{P}_{X_s}[\zeta_{\mathcal{D}} > t - s]] = \mathbb{P}_b[\zeta_{\mathcal{D}} \circ \theta_s > t - s]. \tag{2.22}$$

As $s \to 0$, the probability on the right tends to $\mathbb{P}_b[\zeta_{\mathcal{D}} > t]$, which means nothing but

$$\mathbb{P}_x[\zeta_{\mathcal{D}} > t] \longrightarrow 0, \tag{2.23}$$

as $x \to b$. Thus $\mathbb{P}_x \circ \zeta_{\mathcal{D}}^{-1} \Rightarrow \delta_0$, where $\Rightarrow$ denotes weak convergence. Since also $\mathbb{P}_x \Rightarrow \mathbb{P}_b$ in $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, Theorem 4.28 in Kallenberg [41] yields that

$$\mathbb{P}_x \circ (X, \zeta_{\mathcal{D}})^{-1} \Longrightarrow \mathbb{P}_b \circ (X, 0)^{-1}, \tag{2.24}$$

in $\mathcal{C}(\mathbb{R}_+, \mathbb{R}) \times [0, \infty)$ as $x \to b$. By the continuity of the mapping $(\omega, t) \to \omega_t$ it follows that

$$\mathbb{P}_x \circ X_{\zeta_{\mathcal{D}}}^{-1} \Longrightarrow \mathbb{P}_b \circ X_0^{-1} = \delta_b \tag{2.25}$$

and in particular

$$\mathbb{P}_x \circ (X_{\zeta_{\mathcal{D}}}, \zeta_{\mathcal{D}})^{-1} \Longrightarrow \delta_{b,0}, \tag{2.26}$$

as $x \to b$. By the boundedness and the continuity of the transition density $p(t, x, y)$ it is now clear from (2.18), that $p^{\mathcal{D}}(t, x, y) \to 0$ as $x \to b$. □

We state a regularity result for the transition density $p^{\mathcal{D}}$ together with estimates for $p^{\mathcal{D}}$ and its derivatives. The results can be found in the book of Stroock [68].

Theorem 2.3.0.16. *Suppose that μ and σ belong to $\mathcal{C}_b^\infty(\mathbb{R}, \mathbb{R})$. In addition, let us assume that all derivatives of μ and σ are uniformly bounded, and that σ is uniformly elliptic. Then the transition density $p^{\mathcal{D}} : (0, \infty) \times \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ belongs to $\mathcal{C}^\infty((0, \infty) \times \mathcal{D} \times \mathcal{D}, \mathbb{R})$. For each $n \in \mathbb{N}$, there is a constant Λ_n, depending only on μ and σ and their derivatives through order $n+2$ and $n+3$, respectively, such that the following estimates hold*

$$\begin{aligned} \left| \frac{\partial^m}{\partial t^m} \frac{\partial^\alpha}{\partial x^\alpha} \frac{\partial^\beta}{\partial y^\beta} p^{\mathcal{D}}(t, x, y) \right| \leq & \frac{\Lambda_n e^{\Lambda_n t}}{(t^{1/2} \wedge |x - \mathcal{D}^c| \wedge |y - \mathcal{D}^c|)^{2m+\alpha+\beta+n}} \\ & \times \exp\left(-\frac{(|y-x|^2 \wedge (|x - \mathcal{D}^c| + |x - \mathcal{D}^c|^2))^2}{\Lambda_n t} \right) \end{aligned} \tag{2.27}$$

for all $m, \alpha, \beta \in \mathbb{N}_0$ with $2m + \alpha + \beta \leq n$. Furthermore,

$$\frac{\partial^m}{\partial t^m} p^{\mathcal{D}}(t, x, y) = (\mathcal{A}^m p^{\mathcal{D}}(t, \cdot, y))(x), \tag{2.28}$$

where $\mathcal{A}$ denotes the infinitesimal generator (2.11) *of X.*

Proof. The proof can be found in Chapter 5.2.2 in [68]. □

Remark 2.3.0.17. In the latter theorem relatively strong assumptions on the differentiability of the coefficients μ and σ have been made. The proof for the differentiability relies on an analysis of both the semigroup associated with the operator

$$\mathcal{A} f(x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2} f(x) + \mu(x)\frac{\partial}{\partial x} f(x) \tag{2.29}$$

and on the semigroup associated with the formal adjoint $\mathcal{A}^*$ of $\mathcal{A}$ given by

$$\mathcal{A}^* f(y) = \frac{1}{2}\frac{\partial^2}{\partial y^2}\left((\sigma^2(y)f(y)\right) - \frac{\partial}{\partial y}\left(\mu(y)f(y)\right). \tag{2.30}$$

It is intuitively clear that it suffices to postulate that μ belongs to $\mathcal{C}_b^{n+2}$ and σ belongs to $\mathcal{C}_b^{n+3}$ in order to guarantee that $\frac{\partial^m}{\partial t^m}\frac{\partial^\alpha}{\partial x^\alpha}\frac{\partial^\beta}{\partial y^\beta}p^{\mathcal{D}}(t,x,y)$ exists for $2m+\alpha+\beta \leq n$. And clearly, in this situation the above estimates remain also valid. In order to state differentiability with respect to (t,x), even weaker assumptions about μ and σ are sufficient. Particularly, if Assumption 2.3.0.12 is satisfied, then the functions $\frac{\partial^m}{\partial t^m}\frac{\partial^\alpha}{\partial x^\alpha}p^{\mathcal{D}}(t,x,y)$ exist and are continuous for $2m+\alpha \leq 2$. See the book of Ladyshenskaja et al. [47] for additional information.

Let us state an interesting result concerning the representation of the derivatives of $p^{\mathcal{D}}$ with respect to the forward variable y. This result can also be found in [68].

Theorem 2.3.0.18. *Suppose that the coefficients of the stochastic differential equation* (2.1) *are infinitely many times differentiable with bounded derivatives of all orders and that σ is uniformly elliptic. Then the transition density $p^{\mathcal{D}}(t,x,y)$ exists and is a smooth function on $(0,\infty)\times\mathcal{D}\times\mathcal{D}$. In fact the derivatives of $p^{\mathcal{D}}$ are given by*

$$\frac{\partial^\alpha}{\partial y^\alpha}p^{\mathcal{D}}(t,x,y) = \frac{\partial^\alpha}{\partial y^\alpha}p(t,x,y) - \mathbb{E}_x\left[\frac{\partial^\alpha}{\partial y^\alpha}p(t-\zeta_{\mathcal{D}}, X_{\zeta_{\mathcal{D}}}, y),\ \zeta_{\mathcal{D}} < t\right], \tag{2.31}$$

for every integer $\alpha \in \mathbb{N}$. The stopping time ζ_D is defined by (2.13). *Moreover, for each $n\in\mathbb{N}$, there is a positive constant Λ_n, with the same dependance as the one in Theorem 2.3.0.16, such that for all $\alpha \leq n$,*

$$\sup_{(t,x)\in(0,T]\times\partial\mathcal{D}}\left|\frac{\partial^\alpha}{\partial y^\alpha}p^{\mathcal{D}}(t,x,y)\right| \leq \frac{\Lambda_n}{|y-\mathcal{D}^c|^{\alpha+1}}\exp\left(\Lambda_n T - \frac{|y-\mathcal{D}^c|^2}{\Lambda_n T}\right). \tag{2.32}$$

Proof. Again, the proof can be found in Chapter 5.2.2 in [68]. □

Remark 2.3.0.19. The estimates in formulae (2.27) and (2.32) blow up near the boundary $\partial\mathcal{D}$. This is why they are usually referred to as *interior estimates*. In the sequel we will not make use of these estimates. We mentioned them for the sake of completeness only.

In Theorem 2.3.0.16 we saw that, imposing suitable regularity assumption with respect to the coefficients μ and σ, one is able to show that the function $p^{\mathcal{D}}$ is smooth on $(0,\infty)\times\mathcal{D}\times\mathcal{D}$ and that it satisfies the following differential equation in the backward variable x

$$\frac{\partial}{\partial t}p^{\mathcal{D}}(t,x,y) = [\mathcal{A}\, p^{\mathcal{D}}(t,\cdot,y)](x), \tag{2.33}$$

which is sometimes called *Kolmogorov's backward equation.* As we have stated before, a proof for formula (2.33) is given in Stroock [68]. Here, we give an alternative proof that makes use of very intuitive methods. But before we do so, we have to state an auxiliary result.

Lemma 2.3.0.20. *Assume that μ and σ are continuous and such that the stochastic differential equation* (2.1) *has a (in distribution) unique weak solution for any deterministic initial value. Let $\mathcal{D}$ be a bounded open interval. For any $f \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R})$ set $u(t,x) = E_x[f(X_t),\ t < \zeta_{\mathcal{D}}]$. Then u is a solution to the parabolic partial differential equation*

$$\frac{\partial}{\partial t} u(t,x) = \mathcal{A}u(t,x), \quad \forall x \in \mathcal{D},\, t \geq 0, \tag{2.34}$$

with the initial condition

$$u(0,x) = f(x), \quad \forall x \in \bar{\mathcal{D}} \tag{2.35}$$

and the boundary condition

$$u(t,x) = 0, \quad \forall x \in \partial\mathcal{D},\, t \geq 0. \tag{2.36}$$

Proof. The initial condition follows immediately from the definition of u. Moreover, since f has compact support on $\mathcal{D}$ we conclude that

$$\mathbb{E}_x[f(X_t),\, t < \zeta_{\mathcal{D}}] = \mathbb{E}_x[f(X_{\zeta_{\mathcal{D}} \wedge t})]. \tag{2.37}$$

Dynkin's formula (2.12) applies to the stopping time $\zeta_{\mathcal{D}} \wedge t$. It is obvious that for a function f having compact support on the interval $\mathcal{D}$ we have

$$\int_0^{t \wedge \zeta_{\mathcal{D}}} \mathcal{A}f(X_s)ds = \int_0^t \mathcal{A}f(X_{\zeta_{\mathcal{D}} \wedge s})ds. \tag{2.38}$$

Moreover, the proof of Theorem 2.3.0.15 shows that for a continuous function $f : \bar{\mathcal{D}} \to \mathbb{R}$ and for any $b \in \partial\mathcal{D}$

$$\lim_{\substack{x \to b \\ x \in \mathcal{D}}} \mathbb{E}_x\left[f(X_t),\, t < \zeta_{\mathcal{D}}\right] = 0. \tag{2.39}$$

Combining (2.38) and (2.39) we find that, for all $f \in \mathcal{C}_c^2(\mathcal{D}, \mathbb{R})$ and for all $x \in \bar{\mathcal{D}}$,

$$u(t,x) = \mathbb{E}_x[f(X_t),\, t < \zeta_{\mathcal{D}}] = \mathbb{E}_x[f(X_{\zeta_{\mathcal{D}} \wedge t})] = f(x) + \mathbb{E}_x\left[\int_0^t \mathcal{A}f(X_{\zeta_{\mathcal{D}} \wedge s})ds\right]. \tag{2.40}$$

Since μ and σ are continuous, so is $\mathbb{E}_x[\mathcal{A}f(X_{\zeta_{\mathcal{D}} \wedge s})]$. Hence, by the Fubini-Tonelli theorem, u is continuously differentiable, with respect to t and satisfies $\frac{\partial}{\partial t} u(x,t) = \mathbb{E}_x[\mathcal{A}f(X_{\zeta_{\mathcal{D}} \wedge t})]$. On the other hand one obtains, by the strong Markov property of X and by the fact that f is compactly supported on $\mathcal{D}$, that for all $t, s > 0$,

$$\mathbb{E}_x[u(t, X_{\zeta_{\mathcal{D}}\wedge s})] = \mathbb{E}_x[\mathbb{E}_{X_{\zeta_{\mathcal{D}}\wedge s}}[f(X_{\zeta_{\mathcal{D}}\wedge t})]] = \mathbb{E}_x[f(X_{\zeta_{\mathcal{D}}\wedge(t+s)})] = u(t+s, x). \tag{2.41}$$

For fixed $t > 0$ we infer that the left hand side of

$$\frac{u(t+s,x) - u(t,x)}{s} = \frac{\mathbb{E}_x[u(t, X_{\zeta_{\mathcal{D}}\wedge s})] - u(t,x)}{s} \tag{2.42}$$

converges to $\partial u/\partial t$, as $s \to 0$, and therefore so does the right-hand side. Consequently, u lies in the domain $\mathrm{dom}(\mathcal{A})$ and the assertion follows. □

We obtain the following corollary which gives a concrete characterization for $p^{\mathcal{D}}$ in terms of partial differential equations.

Corollary 2.3.0.21. *Let $\mathcal{D}$ be a bounded open interval. If the transition density $p^{\mathcal{D}}(t,x,y)$ exists, belongs to $\mathcal{C}^2(\mathcal{D},\mathbb{R}) \cap \mathcal{C}(\bar{\mathcal{D}},\mathbb{R})$ with respect to x and is continuously differentiable with respect to t, then for all $y \in \mathcal{D}$ the function $p^{\mathcal{D}}(t,x,y)$ solves the backward Kolmogorov equation*

$$\frac{\partial}{\partial t} u(t,x) = (\mathcal{A}u(t,\cdot))(x), \quad \forall x \in \mathcal{D}, t \geq 0, \tag{2.43}$$

with initial condition

$$u(0,x) = \delta(x-y), \tag{2.44}$$

and with the boundary condition

$$u(t,x) = 0, \quad \forall x \in \partial\mathcal{D},\ t \geq 0. \tag{2.45}$$

In other words, for fixed y the transition density is the fundamental solution of this parabolic partial differential equation with Dirichlet boundary conditions.

Proof. Writing the identity in the preceding Lemma 2.3.0.20 in terms of $p^{\mathcal{D}}$, we obtain for any $f \in \mathcal{C}_c^2(\mathcal{D},\, \mathbb{R})$

$$\frac{\partial}{\partial t} \int_{\mathcal{D}} f(y) p^{\mathcal{D}}(t,x,y) dy = \mathcal{A}\left(\int_{\mathcal{D}} f(y) p^{\mathcal{D}}(t,x,y) dy \right). \tag{2.46}$$

By the compact support of f and the smoothness properties of $p^{\mathcal{D}}$, we may interchange integration and differentiation on both sides. From the equation

$$\int_{\mathcal{D}} \left\{ \frac{\partial}{\partial t} - \mathcal{A} \right\} p^{\mathcal{D}}(t,x,y) f(y) dy = 0, \tag{2.47}$$

which holds for any test function f, we then conclude by a continuity argument. It remains to show (2.45). But this boundary condition follows directly from Theorem 2.3.0.15. □

Moreover, we can infer the following corollary which states that the transition density of the killed process satisfies the so-called *Kolmogorov forward equation.* Note that sometimes this equation is also called the *Fokker-Planck equation.*

Corollary 2.3.0.22. *If the transition density $p^{\mathcal{D}}(t,x,y)$ exists, belongs to $\mathcal{C}^2(\mathcal{D},\mathbb{R})\cap\mathcal{C}(\bar{\mathcal{D}},\mathbb{R})$ with respect to y and is continuously differentiable with respect to t, then $p^{\mathcal{D}}(t,x,y)$ solves, for all $x\in\mathcal{D}$, the forward Kolmogorov equation*

$$\frac{\partial u}{\partial t}(y,t)=(\mathcal{A}^*u(\cdot,t))(y),\quad \forall y\in\mathcal{D},t\geq 0, \tag{2.48}$$

with initial condition

$$u(y,0)=\delta(y-x), \tag{2.49}$$

and with the boundary condition

$$u(t,y)=0,\quad \forall y\in\partial\mathcal{D},\ t\geq 0, \tag{2.50}$$

where the operator

$$\mathcal{A}^*f(y)=\frac{1}{2}\frac{\partial^2}{\partial y^2}\Big((\sigma^2(y)f(y)\Big)-\frac{\partial}{\partial y}\left(\mu(y)f(y)\right) \tag{2.51}$$

denotes the formal adjoint of $\mathcal{A}$. Hence, for fixed x, the transition density is the fundamental solution of the parabolic partial differential equation with the adjoint operator $\mathcal{A}^*$ and Dirichlet boundary conditions.

Proof. For $f\in\mathcal{C}^2_c(\mathcal{D},\mathbb{R})$ let us evaluate $\mathbb{E}_x[\mathcal{A}f(X_{\zeta_{\mathcal{D}}\wedge t})]$ in two different ways. First, we obtain by definition

$$\mathbb{E}_x[\mathcal{A}f(X_{\zeta_{\mathcal{D}}\wedge t}))]=\int_{\mathcal{D}}\mathcal{A}f(y)p^{\mathcal{D}}(t,x,y)dy=\int_{\mathcal{D}}f(y)(\mathcal{A}^*p^{\mathcal{D}}(t,x,\cdot))(y)dy. \tag{2.52}$$

On the other hand, by dominated convergence and by Dynkin's formula we find

$$\int_{\mathcal{D}}f(y)\frac{\partial}{\partial t}p^{\mathcal{D}}(t,x,y)dy=\frac{\partial}{\partial t}\mathbb{E}_x[f(X_{\zeta_{\mathcal{D}}\wedge t})]=\mathbb{E}_x[\mathcal{A}f(X_{\zeta_{\mathcal{D}}\wedge t})]. \tag{2.53}$$

We conclude again by testing this identity with all $f\in\mathcal{C}^2_c(\mathcal{D},\ \mathbb{R})$. Finally, the boundary condition follows directly from the definition of a killed diffusion or from (2.16). $\square$

Of course, the Corollaries 2.3.0.21 and 2.3.0.22 alone would not be satisfactory, for we had to impose the existence of a transition density with restrictive regularity properties. But together with the existence and smoothness results concerning $p^{\mathcal{D}}$ in Theorem 2.3.0.16 and Theorem 2.3.0.18 they become quite handy. Also see Remark 2.3.0.17. Overall, we found sufficient conditions for the existence of a fundamental solution to both the backward and the forward Kolmogorov equation with Dirichlet boundary conditions, which are described by (2.43), (2.44), (2.45) and (2.48), (2.49), (2.50), respectively.

Conversely, a fundamental solution to the backward Kolmogorov equation with Dirichlet boundary conditions coincides with the transition density of the corresponding killed diffusion. This follows directly from the next proposition and the subsequent corollary.

Proposition 2.3.0.23. *Let us consider a bounded open interval $\mathcal{D} \subset \mathbb{R}$ and assume that $f : \bar{\mathcal{D}} \to \mathbb{R}$ is continuous. Assume that μ and σ are such that the stochastic differential equation* (2.1) *has a (in distribution) unique weak solution for any deterministic initial value. If $u \in \mathcal{C}^{1,2}((0,\infty)\times\mathcal{D},\mathbb{R})\cap\mathcal{C}(\mathbb{R}_+\times\bar{\mathcal{D}},\mathbb{R})$ denotes a solution to the Cauchy problem*

$$\frac{\partial}{\partial t}u(t,x) = \mathcal{A}u(t,\cdot)(x), \tag{2.54}$$

with initial condition

$$u(0,x) = f(x), \quad \text{for } x \in \mathcal{D}, \tag{2.55}$$

and boundary condition

$$u(t,x) = 0, \quad \text{for } x \in \partial\mathcal{D}, \text{ and for all } \ t \geq 0, \tag{2.56}$$

then u is given by

$$u(t,x) = \mathbb{E}_x\big[f(X_t)\mathbb{1}_{\{t<\tau_{\mathcal{D}}\}}\big]. \tag{2.57}$$

Proof. Let u denote a solution to (2.54), with initial condition (2.55) and boundary condition (2.56). Let $t > 0$ be fixed. We apply Itô's formula to the process

$$u(t-s, X_s), \quad 0 \leq s \leq t, \tag{2.58}$$

and we integrate from 0 to $\zeta_{\mathcal{D}} \wedge t$ in order to obtain

$$\begin{aligned} u(t-\zeta_{\mathcal{D}}\wedge t, X_{\zeta_{\mathcal{D}}\wedge t}) = u(t,x) &- \int_0^{\zeta_{\mathcal{D}}\wedge t} \frac{\partial}{\partial a}u(t-a,X_s)\big|_{a=s}ds \\ &+ \int_0^{\zeta_{\mathcal{D}}\wedge t} \mathcal{A}u(t-s,\cdot)(X_s)ds \\ &+ \int_0^{\zeta_{\mathcal{D}}\wedge t} \frac{\partial}{\partial y}u(t-s,y)\big|_{y=X_s}\sigma(X_s)dB_s. \end{aligned} \tag{2.59}$$

Since u solves (2.54), taking expectations on both sides yields the equation

$$\begin{aligned} u(t,x) &= \mathbb{E}_x[u(t-\zeta_{\mathcal{D}}\wedge t, X_{\zeta_{\mathcal{D}}\wedge t})] \\ &= \mathbb{E}_x\big[u(0,X_t)\mathbb{1}_{\{t<\zeta_{\mathcal{D}}\}}\big] + \mathbb{E}_x\big[u(t-\zeta_{\mathcal{D}}, X_{\zeta_{\mathcal{D}}})\mathbb{1}_{\{t\geq\zeta_{\mathcal{D}}\}}\big]. \end{aligned} \tag{2.60}$$

The boundary condition (2.56) implies that the second term on the right hand side of equation (2.60) vanishes. By the initial condition (2.55) the first expression on the right hand side of (2.60) coincides with the right hand side of (2.57). □

From the latter proposition we are able to infer the following uniqueness result.

Corollary 2.3.0.24. *Let $\mathcal{D} \subset \mathbb{R}$ be a bounded open interval. Consider a function $p^{\mathcal{D}} \in \mathcal{C}^{1,2,0}((0,\infty)\times\mathcal{D}\times\mathcal{D},\mathbb{R})\cap\mathcal{C}(\mathbb{R}_+\times\bar{\mathcal{D}}\times\mathcal{D},\mathbb{R})$ and let us assume that, for all $y \in \mathcal{D}$, $p^{\mathcal{D}}(t,x,y)$ is a fundamental solution to the initial/boundary value problem given by* (2.43) - (2.45)*. Then $p^{\mathcal{D}}$ is unique.*

Proof. Let $y \in \mathcal{D}$ arbitrary but fixed. Consider a series of smooth functions $\delta^{(n)} : \mathcal{D} \to \mathbb{R}$, having compact support on $\mathcal{D}$, such that

$$\lim_{n\to\infty} \delta^{(n)} = \delta_y \tag{2.61}$$

in the weak*-topology of the dual space of $\mathbb{R}$. Here δ_y denotes the Dirac measure of y. By the smoothness of $\delta^{(n)}$, the function

$$(t,x) \longmapsto \int_{\mathcal{D}} \delta^{(n)}(z)p^{\mathcal{D}}(t,x,z)dz \tag{2.62}$$

is a solution to (2.54), with boundary condition

$$\int_{\mathcal{D}} \delta^{(n)}(z)p^{\mathcal{D}}(t,x,z)dz = 0, \quad \text{for } x \in \partial\mathcal{D}, \text{ and for all } t \geq 0, \tag{2.63}$$

and with initial condition

$$\int_{\mathcal{D}} \delta^{(n)}(z)p^{\mathcal{D}}(0,x,z)dz = \delta^{(n)}(x), \quad \forall x \in \mathcal{D}. \tag{2.64}$$

Assume that there is another solution $\bar{p}^{\mathcal{D}}$ to (2.43), (2.45). Then the function

$$(t,x) \longmapsto \int_{\mathcal{D}} \delta^{(n)}(z)\bar{p}^{\mathcal{D}}(t,x,z)dy \tag{2.65}$$

also solves (2.54) with boundary condition (2.63) and initial condition (2.64). By Proposition 2.3.0.23 we find that

$$\int_{\mathcal{D}} \delta^{(n)}(z)p^{\mathcal{D}}(t,x,z)dz = \int_{\mathcal{D}} \delta^{(n)}(z)\bar{p}^{\mathcal{D}}(t,x,z)dz = \mathbb{E}_x[\delta^{(n)}(X_t)\mathbb{1}_{\{t<\tau_{\mathcal{D}}\}}], \tag{2.66}$$

for all $n \in \mathbb{N}$. We conclude by letting $n \to \infty$. Consequently, the functions $p^{\mathcal{D}}$ and $\bar{p}^{\mathcal{D}}$ coincide on the interior of $\mathcal{D}$. Equality at the boundary follows from the boundary condition. This completes the proof of the uniqueness result. □

3 Joint Densities of the Maximum and the Minimum of Diffusion Processes

3.1 Introduction

In Section 3.2 we derive an existence result for the joint density of one-dimensional diffusions and their running maximum and running minimum. In Section 3.3 different ways to calculate the joint probabilities and the joint densities of the triplet (H_t, L_t, X_t) are depicted. Also, some results concerning the Laplace transform of hitting time densities are presented. Finally, in Section 3.4, we consider some important examples.

3.2 Existence of joint densities

For a class of one-dimensional diffusion processes X we prove an existence result for the densities of (H, L) and (H, L, X), respectively. As usual, H denotes the running maximum and L denotes the running minimum of X. Malliavin calculus is necessary to prove our statements. Therefore we introduce some elementary concepts concerning this topic first. We restrict ourselves to a presentation of the facts that are necessary for our purposes. The notations of this section are oriented towards the notations in the book of Nualart [50]. They slightly deviate from the notations of Chapter 2.

3.2.1 The Malliavin calculus - a brief introduction

The results presented in this paragraph are special cases of the results outlined in the book of Nualart [50].

Let $\mathcal{H}$ be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, the corresponding norm will be denoted with $\| \cdot \|_{\mathcal{H}}$. Let $W = \{W(h),\ h \in \mathcal{H}\}$ be an isonormal Gaussian process indexed by the elements of $\mathcal{H}$ on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where the σ-field $\mathcal{F}$ is supposed to be generated by W. We want to introduce the derivative DF of a square integrable random variable $F : \Omega \to \mathbb{R}$. This means that we intend to differentiate F with respect to the chance parameter $\omega \in \Omega$. In usual applications, the space Ω is the space of continuous functions $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$. We will discuss this special case in the next section, when we are going to consider stochastic differential equations. But let us come back to the introduction. We denote by $\mathcal{C}_p^\infty(\mathbb{R}^n)$ the set of all infinitely many times continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$ such that f and all of its partial derivatives have polynomial growth. The Schwartz-space $\mathcal{S}$ is the class of smooth random variables of the form

$$F = f(W(h_1), \dots, W(h_n)), \tag{3.1}$$

where $f \in \mathcal{C}_p^\infty(\mathbb{R}^n)$, $h_1, \dots, h_n \in \mathcal{H}$ and $n \in \mathbb{N}$.

Definition 3.2.1.1. The derivative of a smooth random variable $F \in \mathcal{S}$ of the form (3.1) is the $\mathcal{H}$-valued random variable

$$DF = \sum_{i=1}^{n} \partial_i f(W(h_1), \dots, W(h_n)) h_i, \tag{3.2}$$

where $\partial_i f = \frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$.

For example, we have $DW(h) = h$. Moreover, DF can be interpreted as a directional derivative, since for any element $h \in \mathcal{H}$ we have

$$\begin{aligned} \langle DF, h \rangle_{\mathcal{H}} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \Big[& f(W(h_1) + \epsilon \langle h_1, h \rangle_{\mathcal{H}}, \dots, W(h_n) + \epsilon \langle h_1, h \rangle_{\mathcal{H}}) \\ & - f(W(h_1), \dots, W(h_n)) \Big]. \end{aligned} \tag{3.3}$$

Roughly speaking, the scalar product $\langle DF, h \rangle_{\mathcal{H}}$ is the derivative at $\epsilon = 0$ of the random variable F composed with the shifted process $\{W(g) + \epsilon \langle g, h \rangle_{\mathcal{H}},\ g \in \mathcal{H}\}$.

Theorem 3.2.1.2. *The operator D is closable from $L^p(\Omega)$ to $L^p(\Omega; \mathcal{H})$ for any $p \geq 1$.*

Proof. The proof is based on the integration-by-parts formula

$$\mathbb{E}[G \langle DF, h \rangle_{\mathcal{H}}] = \mathbb{E}[-F \langle DG, h \rangle_{\mathcal{H}} + FGW(h)], \tag{3.4}$$

that is valid for $F, G \in \mathcal{S}$ and $h \in \mathcal{H}$. Proofs for this formula and the above theorem can be found in [50], Chapter 1.2, page 25 ff. □

Definition 3.2.1.3. For any $p \geq 1$ we will denote the domain of the operator D in $L^p(\Omega)$ by $\mathbb{D}^{1,p}$. This means that $\mathbb{D}^{1,p}$ is the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} := \Big(\mathbb{E}[|F|^p] + \mathbb{E}[\|DF\|_{\mathcal{H}}^p] \Big)^{1/p}. \tag{3.5}$$

Remark 3.2.1.4. For $p = 2$, the space $\mathbb{D}^{1,2}$ is a Hilbert space with the scalar product

$$\langle F, G \rangle_{1,2} = \mathbb{E}[FG] + \mathbb{E}[\langle DF, DG \rangle_{\mathcal{H}}]. \tag{3.6}$$

Furthermore, an iteration of the operator D can be defined in such a way that for a random variable $F \in \mathcal{S}$, the iterated derivative D^k is a random variable with values in

$\mathcal{H}^{\otimes k}$. Therefore it is necessary to introduce the following seminorms on $\mathcal{S}$ defined by

$$\|F\|_{k,p} := \Big(\mathbb{E}[|F|^p] + \sum_{j=1}^{k} \mathbb{E}[\|D^j F\|_{\mathcal{H}^{\otimes j}}^p]\Big)^{1/p}, \quad p \geq 1,\ k \in \mathbb{N}. \tag{3.7}$$

Then the space $\mathbb{D}^{k,p}$ is defined as the closure of $\mathcal{S}$ in $L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$. We omit a further discussion of the spaces $\mathbb{D}^{k,p}$, since in the sequel we will only work on the spaces $\mathbb{D}^{1,2}$ or $\mathbb{D}^{1,p}$, respectively.

We end this introduction by stating a sufficient criterion for a random vector to have a Lebesgue-density.

Theorem 3.2.1.5. *Let $F = (F_1, \dots, F_m)$ be a random vector that satisfies the following two conditions:*

(i) F_i belongs to the space $\mathbb{D}^{1,p}$, for a $p > 1$ and for all $i = 1, \dots, m$.

(ii) The matrix $M_F = \big(\langle DF_i, DF_j\rangle_{\mathcal{H}}\big)_{i,j=1,\dots,m}$ is invertible a.s.

Then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.

Proof. The proof can be found in Nualart [50], Chapter 2.1, p. 97 ff. □

3.2.2 Malliavin calculus and stochastic differential equations

Now suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is the canonical probability space associated with a one-dimensional Brownian motion $(W_t,\ t \in [0,T])$ on a finite interval $[0,T]$. As in Chapter 2, by the canonical probability space, we mean $\Omega = \mathcal{C}([0,T],\mathbb{R})$. Here, $\mathbb{P}$ denotes the one-dimensional Wiener measure and $\mathcal{F}$ denotes the completion (with respect to $\mathbb{P}$) of the Borel-σ-field generated by the topology of uniform convergence. The underlying Hilbert space in the present situation is $\mathcal{H} = L^2([0,T],\mathbb{R})$ and the scalar product on $\mathcal{H}$ will be denoted with $\langle\cdot,\cdot\rangle_{L^2}$.

Let the coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$, be measurable functions that satisfy a global Lipschitz condition with linear growth.

Condition 3.2.2.1. *We assume that there is a constant $K > 0$ such that*

$$|\sigma(x) - \sigma(y)| + |\mu(x) - \mu(y)| \leq K|x-y|, \tag{3.8}$$

for any $x, y \in \mathbb{R}$. In addition, we assume that σ is uniformly elliptic.

Condition 3.2.2.1 ensures, for all $x \in \mathbb{R}$, the existence of a unique continuous solution $X = \{X_t,\ t \in [0,T]\}$ to the stochastic differential equation

$$X_t = x + \int_0^t \sigma(X_s)dW_s + \int_0^t \mu(X_s)ds, \tag{3.9}$$

such that for all $t \in [0,T]$ the random variable X_t belongs to the space $\mathbb{D}^{1,2}$ for all $p \geq 2$.

Some general properties of the Malliavin derivatives of X are stated in the following theorem.

Theorem 3.2.2.2. *Suppose Condition 3.2.2.1 is satisfied and, for $x \in \mathbb{R}^m$, let $X = \{X_t,\ t \in [0,T]\}$ be the solution to* (3.9). *Then X_t belongs to $\mathbb{D}^{1,\infty}$ for any $t \in [0,T]$. Moreover,*

$$\sup_{0 \leq r \leq t} \mathbb{E}\left(\sup_{r \leq s \leq T} |D_r X_s|^p \right) < \infty, \tag{3.10}$$

and the derivative $D_r X_t$ satisfies the following linear equation

$$D_r X_t = \sigma(X_r) + \int_r^t \bar{\sigma}(s)(D_r X_s) dW_s + \int_r^t \bar{\mu}(s)(D_r X_s) ds \tag{3.11}$$

for $r \leq t$ a.e.,

$$D_r X(t) = 0 \tag{3.12}$$

for $r > t$ a.e., and the $\bar{\sigma}(s)$ and $\bar{\mu}(s)$ are uniformly bounded and adapted one-dimensional processes.

Furthermore, if the coefficients μ and σ of equation (3.9) *are continuously differentiable, then one can write*

$$\bar{\sigma}(s) = \sigma'(X_s) \tag{3.13}$$

and

$$\bar{\mu}(s) = \mu'(X_s). \tag{3.14}$$

Proof. See e.g. Nualart [50], Theorem 2.2.1, page 119 ff. □

The next lemma yields a suitable representation of the process DX_t.

Lemma 3.2.2.3. *Let σ and μ be continuously differentiable functions on $\mathbb{R}$ with uniformly bounded first derivatives and let σ be uniformly elliptic. The Malliavin derivative of the solution to the stochastic differential equation*

$$X_t = X_0 + \int_0^t \sigma(X_s) dW_s + \int_0^t \mu(X_s) ds, \quad t \in [0,T], \tag{3.15}$$

satisfies

$$D_r X_t = \sigma(X_r) \mathbb{1}_{\{r \leq t\}} \exp\left(\int_r^t \sigma'(X_s) dW_s + \int_r^t [\mu' - \frac{1}{2}(\sigma')^2](X_s) ds \right), \quad r,t \in [0,T]. \tag{3.16}$$

Proof. The assumptions about σ and μ guarantee the existence of a unique strong solution to (3.15). By Theorem 3.2.2.2 and for $r \leq t$ the process $D_r X_t$ must satisfy the equation

$$D_r X_t = \sigma(X_r) + \int_r^t \sigma'(X_s) D_r X_s dW_s + \int_r^t \mu'(X_s) D_r X_s ds. \tag{3.17}$$

In the sequel let $r \leq t$. We apply Itô's formula to

$$Z_t = \sigma(X_r) \exp(Y_t), \tag{3.18}$$

where

$$Y_t = \int_r^t \sigma'(X_s) dW_s + \int_r^t \left[\mu' - \frac{1}{2}(\sigma')^2\right](X_s) ds, \tag{3.19}$$

in order to obtain

$$dZ_t = \sigma(X_r) \exp(Y_t) dY_t + \frac{1}{2}\sigma(X_r) \exp(Y_t) d\langle Y, Y\rangle_t \,. \tag{3.20}$$

The quadratic variation of Y is

$$\langle Y, Y\rangle_t = \int_r^t (\sigma')^2(X_s)\, ds. \tag{3.21}$$

Hence, equation (3.20) becomes

$$\begin{aligned} dZ_t &= \sigma'(X_t)\sigma(X_r) \exp(Y_t) dW_t + \sigma(X_r) \exp(Y_t)[\mu' - \frac{1}{2}(\sigma')^2](X_t) dt \\ &\quad + \frac{1}{2}(\sigma')^2(X_t)\sigma(X_r) \exp(Y_t) dt \qquad (3.22) \\ &= \sigma'(X_t)\sigma(X_r) \exp(Y_t) dW_t + \mu'(X_t)\sigma(X_r) \exp(Y_t) dt \\ &= \sigma'(X_t) Z_t dW_t + \mu'(X_t) Z_t dt. \qquad (3.23) \end{aligned}$$

Thus we have shown that Z_t satisfies (3.17). Since $Z_r = \sigma(X_r)$, this yields the assertion. □

The following lemma will be a helpful means to derive a representation of the Malliavin derivative of the running maximum of a diffusion process.

Lemma 3.2.2.4. *Let $(X_t,\, t \in I)$ be a continuous process parametrized by a compact interval I. Suppose that*

(i) $\mathbb{E}[\sup_{t\in I} X_t^2] < \infty$,

(ii) for any $t \in I$, $X_t \in \mathbb{D}^{1,2}$, the $\mathbb{R}$-valued process $\{DX_t, t \in I\}$ possesses a continuous version, and $\mathbb{E}\big[\sup_{t\in I} \|DX_t\|_{L^2}^2\big]$,

then the random variable $H = \sup_{t\in I} X_t$ *belongs to* $\mathbb{D}^{1,2}$ *and* $DH = DX_t$ *a.s. on the set* $\{(t,\omega) : X_t(\omega) = H(\omega)\}$.

Proof. See Nualart [50], p. 110, proof of Proposition 2.1.11. □

Henceforth, let $T > 0$ be fixed and let τ_H denote the random time where the process $(X_t,\, t \in [0,T])$ attains its maximum. Then τ_H is uniquely defined a.s., see e.g. Karatzas and Shreve [43]. Lemma 3.2.2.4 immediately implies that

$$DH = DH_T = DX_{\tau_H} \quad \text{a.s.,} \tag{3.24}$$

where $H = H_T = \sup_{0\le t\le T} X_t$ and with the understanding that DX_{τ_H} is the random variable $\omega \mapsto DX_{\tau_H(\omega)}(\omega)$. In the sequel we will make use of the notation

$$\begin{aligned} D_r H_T &= D_r X_{\tau_H} \\ &=: \sigma(X_r)\mathbb{1}_{\{r\le\tau^H\}} \exp\left(\int_r^{\tau^H} \sigma'(X_s)dW_s - \int_r^{\tau^H} \left[\mu' - \frac{1}{2}(\sigma')^2\right](X_s)ds\right), \end{aligned} \tag{3.25}$$

where the integral

$$\int_r^{\tau^H} \sigma'(X_s)dW_s \tag{3.26}$$

is defined on the set $\{r \le \tau_H\}$, and it is supposed to be the random variable

$$\omega \longmapsto Y(\tau_H(\omega))(\omega), \tag{3.27}$$

with $Y(t) = \int_r^t \sigma'(X_s)dW_s$ for $r < t$.

For the minimum $L = L_T = \inf_{0\le t\le T} X_t$, the Malliavin derivative $DL = DL_T = DX_{\tau_L}$ is defined in an analogous way.

Now that we have established representations for DH and DL, the aim is to calculate the expressions $\langle DH, DH\rangle$, $\langle DH, DL\rangle$ and $\langle DL, DL\rangle$. For $a < b$ we introduce the notation

$$g(a,b) = \exp\left(\int_a^b \sigma'(X_s)dW_s - \int_a^b \left[\mu' - \frac{1}{2}(\sigma')^2\right](X_s)ds\right). \tag{3.28}$$

Let us assume that we are on the set $\{\tau_H < \tau_L\}$. We define

$$g(\tau_H,\tau_L) := \exp\left(\int_{\tau_H}^{\tau_L} \sigma'(X_s)dW_s - \int_{\tau_H}^{\tau_L} \left[\mu' - \frac{1}{2}(\sigma')^2\right](X_s)ds\right), \tag{3.29}$$

where the stochastic integral in the exponent of (3.29), that is the expression

$$\int_{\tau_H}^{\tau_L} \sigma'(X_s)dW_s, \tag{3.30}$$

is supposed to be the random variable

$$\omega \longmapsto \tilde{Y}(\tau_H(\omega), \tau_L(\omega))(\omega), \tag{3.31}$$

with $\tilde{Y}(r,t) = \int_r^t \sigma'(X_s)dW_s$ for $r < t$.

Now, we are in the position to state the following lemma.

Lemma 3.2.2.5. *Suppose we are on the set $\{\tau_H < \tau_L\}$. Then*

$$\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_L)^2 dr = g(\tau_H,\tau_L)^2 \int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \tag{3.32}$$

and

$$\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr = g(\tau_H,\tau_L) \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr. \tag{3.33}$$

Proof. By the definition of $g(\tau_H,\tau_L)$ in (3.29), on the set $\{\tau_H < \tau_L\}$, we have

$$g(r,\tau_L) = g(r,\tau_H)g(\tau_H,\tau_L), \quad 0 \le r \le \tau_H, \tag{3.34}$$

and the result follows. $\square$

Remark 3.2.2.6. Of course, on the set $\{\tau_L < \tau_H\}$, the expression $g(\tau_L,\tau_H)$ can be defined analogously to (3.29). The definitions of the random variables $g(\tau_L,T)$ and $g(\tau_H,T)$ are straightforward, since $\tau_H < T$ and $\tau_L < T$ a.s. For the respective sets, results analogous to Lemma 3.2.2.5 hold.

3.2.3 Applications - Existence of the joint density

In order to prove the existence of a joint density of (H,L) one has to show that the matrix

$$M_{(H,L)} = \begin{pmatrix} \langle DH, DH\rangle_{L^2} & \langle DH, DL\rangle_{L^2} \\ \langle DL, DH\rangle_{L^2} & \langle DL, DL\rangle_{L^2} \end{pmatrix} \tag{3.35}$$

is positive definite a.s. One can likewise show the existence of a joint density of (H,L,X) by considering the Malliavin matrix

$$M_{(H,L,X)} = \begin{pmatrix} \langle DH, DH\rangle_{L^2} & \langle DH, DL\rangle_{L^2} & \langle DH, DX\rangle_{L^2} \\ \langle DL, DH\rangle_{L^2} & \langle DL, DL\rangle_{L^2} & \langle DL, DX\rangle_{L^2} \\ \langle DX, DH\rangle_{L^2} & \langle DX, DL\rangle_{L^2} & \langle DX, DX\rangle_{L^2} \end{pmatrix}. \tag{3.36}$$

Theorem 3.2.3.1. *Let the coefficients $\sigma, \mu : \mathbb{R} \to \mathbb{R}$ be continuously differentiable functions with uniformly bounded first derivatives. Moreover, suppose that σ is uniformly* *bounded away from zero. Then the matrix*

$$M_{(H,L)} = \begin{pmatrix} \langle DH, DH \rangle_{L^2} & \langle DH, DL \rangle_{L^2} \\ \langle DL, DH \rangle_{L^2} & \langle DL, DL \rangle_{L^2} \end{pmatrix} \tag{3.37}$$

is positive definite a.s.

Proof. Clearly, on $\{\tau_H < \tau_L\}$, we have $\tau^H \wedge \tau^L = \tau_H$, and by Lemma 3.2.2.5 we find

$$\begin{aligned}
&\det\left(M_{(H,L)}\right) \\
&= \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr \\
&\qquad - \left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr\right)^2 \\
&= \int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_L)^2 dr + \int_{\tau_H}^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right) \\
&\qquad - \left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr\right)^2 \\
&= \left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right) \\
&\qquad \times \left(g(\tau_H,\tau_L)^2 \int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr + \int_{\tau_H}^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right) \\
&\quad - g(\tau_H,\tau_L)^2 \left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right)^2 \\
&= \int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \int_{\tau_H}^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr.
\end{aligned} \tag{3.38}$$

The expression on the right hand side can easily be shown to be positive a.s. For $\{\tau_L < \tau_H\}$, we proceed analogously and obtain

$$\det\left(M_{(H,L)}\right) = \int_0^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr \int_{\tau_L}^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr. \tag{3.39}$$

Since $\mathbb{P}[\tau_H = \tau_L] = 0$ this yields the positive definiteness of $M_{(H,L)}$. □

Theorem 3.2.3.2. *Again let $\sigma, \mu : \mathbb{R} \to \mathbb{R}$ be continuously differentiable functions with uniformly bounded first derivatives and let σ be uniformly bounded away from zero. Then the Malliavin matrix*

$$M_{(H,L,X)} = \begin{pmatrix} \langle DH, DH \rangle_{L^2} & \langle DH, DL \rangle_{L^2} & \langle DH, DX \rangle_{L^2} \\ \langle DL, DH \rangle_{L^2} & \langle DL, DL \rangle_{L^2} & \langle DL, DX \rangle_{L^2} \\ \langle DX, DH \rangle_{L^2} & \langle DX, DL \rangle_{L^2} & \langle DX, DX \rangle_{L^2} \end{pmatrix} \tag{3.40}$$

is positive definite a.s.

Proof. Let $g(a,b)$ be defined as in (3.28). Then, on the set $\{\tau_H < \tau_L < T\}$, we have

$$\begin{aligned}
&\det\left(M_{(H,L,X)}\right)\\
&= \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \times \Bigg[\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr \int_0^T \sigma(X_r)^2 g(r,T)^2 dr \\
&\qquad - \left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L) g(r,T) dr\right)^2\Bigg] \\
&- \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr \times \Bigg[\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr \int_0^T \sigma(X_r)^2 g(r,T)^2 dr \\
&\qquad - \left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L) g(r,T) dr\right)\left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,T) dr\right)\Bigg] \\
&+ \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,T) dr \times \Bigg[\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr \int_0^{\tau_L} \sigma(X_r)^2 g(r,T) g(r,\tau_L) dr \\
&\qquad - \left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right)\left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,T) dr\right)\Bigg]. \qquad (3.41)
\end{aligned}$$

First, let us consider the last term on the right hand side of formula (3.41) separately. After some calculus and by means of Lemma 3.2.2.5, we find

$$\begin{aligned}
&\left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H) g(r,\tau_L) dr\right)\left(\int_0^{\tau_L} \sigma(X_r)^2 g(r,T) g(r,\tau_L) dr\right) \\
&\qquad - \left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right)\left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H) g(r,T) dr\right) \\
&= g(\tau_H,\tau_L) g(\tau_L,T)\left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right)\left(\int_0^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right) \\
&\qquad - g(\tau_H,T)\left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right)\left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right) \\
&= \left(\int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right)\left(\int_0^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right) \\
&\qquad \times \left[\underbrace{g(\tau_H,\tau_L) g(\tau_L,T)}_{g(\tau_H,T)} - g(\tau_H,T)\right] \\
&= 0, \qquad (3.42)
\end{aligned}$$

and hence the last addend in (3.41) vanishes. Thus, one obtains for $\det\left(M_{(H,L,X)}\right)$:

$$\begin{aligned}
&\det\left(M_{(H,L,X)}\right)\\
&= \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \times \Bigg[\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr \Bigg(g(\tau_L,T)^2 \int_0^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr
\end{aligned}$$

$$
\begin{aligned}
&\qquad + \int_{\tau_L}^{T} \sigma(X_r)^2 g(r,T)^2 dr\bigg) - g(\tau_L,T)^2 \left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right)^2\Bigg] \\
&- \ g(\tau_H,\tau_L) \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \\
&\quad \times \Bigg[g(\tau_H,\tau_L) \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \left(g(\tau_L,T)^2 \int_0^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr + \int_{\tau_L}^{T} \sigma(X_r)^2 g(r,T)^2 dr \right) \\
&\qquad - \ g(\tau_L,T) g(\tau_H,T) \left(\int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr\right) \left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right)\Bigg] \\
&= \int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \int_0^{\tau^L} \sigma(X_r)^2 g(r,\tau_L)^2 dr \int_{\tau_L}^{T} \sigma(X_r)^2 g(r,T)^2 dr \\
&\quad - \ g(\tau_H,\tau_L)^2 \left(\int_0^{\tau^H} \sigma(X_r)^2 g(r,\tau_H)^2 dr\right)^2 \int_{\tau_L}^{T} \sigma(X_r)^2 g(r,T)^2 dr \\
&= \int_{\tau_L}^{T} \sigma(X_r)^2 g(r,T)^2 dr \int_{\tau_H}^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr \int_0^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \,. \qquad (3.43)
\end{aligned}
$$

Analogously, we find that, on the set $\{\tau_L < \tau_H < T\}$,

$$
\begin{aligned}
&\det\left(M_{(H,L,X)}\right) \\
&= \int_{\tau_H}^{T} \sigma(X_r)^2 g(r,T)^2 dr \int_{\tau_L}^{\tau_H} \sigma(X_r)^2 g(r,\tau_H)^2 dr \int_0^{\tau_L} \sigma(X_r)^2 g(r,\tau_L)^2 dr. \qquad (3.44)
\end{aligned}
$$

Since

$$\mathbb{P}[\tau_H = \tau_L] = \mathbb{P}[\tau_H = T] = \mathbb{P}[\tau_L = T] = 0, \qquad (3.45)$$

this concludes the proof of our result. □

Remark 3.2.3.3. If X denotes Brownian motion, we have

$$D_r H_T = \mathbb{1}_{\{r \leq \tau_H\}}, \quad D_r L_T = \mathbb{1}_{\{r \leq \tau_L\}} \quad \text{and} \quad D_r X_T = \mathbb{1}_{\{r \leq T\}}. \qquad (3.46)$$

Therefore, the Malliavin matrices have a very simple form, namely

$$M_{(H,L)} = \begin{pmatrix} \tau_H & \tau_H \wedge \tau_L \\ \tau_H \wedge \tau_L & \tau_L \end{pmatrix} \qquad (3.47)$$

and

$$M_{(H,L,X)} = \begin{pmatrix} \tau_H & \tau_H \wedge \tau_L & \tau_H \\ \tau_H \wedge \tau_L & \tau_L & \tau_L \\ \tau_H & \tau_L & T \end{pmatrix}. \qquad (3.48)$$

Straightforward calculations yield that, on the set $\{\tau_H < \tau_L\}$,

$$\det M_{(H,L)} = \tau_H \tau_L - \tau_H^2 = (\tau_L - \tau_H)\tau_H \tag{3.49}$$

and

$$\det M_{(H,L,X)} = \tau_H(T\tau_L - \tau_L^2) - \tau_H(T\tau_H - \tau_L\tau_H) = (T - \tau_L)(\tau_L - \tau_H)\tau_H. \tag{3.50}$$

For $\{\tau_H < \tau_L\}$ analogously. Let us compare these two formulae to the determinants we found in the proofs of the Theorems 3.2.3.1 and 3.2.3.2. We see that (3.49) and (3.50) are only special cases of (3.38) and (3.43), respectively.

We conclude this section by stating the following theorem.

Theorem 3.2.3.4. *Let σ and μ be continuously differentiable functions on $\mathbb{R}$ with uniformly bounded first derivatives and let σ be uniformly elliptic. Let $H = H_T = \sup_{0 \le t \le T} X_t$ and $L = L_T = \inf_{0 \le t \le T} X_t$, where the process X is the solution to the stochastic differential equation*

$$X_t = X_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t \mu(X_s)ds, \quad X_0 = x, \tag{3.51}$$

on the interval $[0, T]$. Then the random vectors (H_T, L_T) and (H_T, L_T, X_T) are absolutely continuous with respect to the Lebesgue measure of $\mathbb{R}^2$ or of $\mathbb{R}^3$, respectively.

Proof. Since the random variables H_T, L_T and X_T belong to $\mathbb{D}^{1,2}$ (see Lemma 3.2.2.4), the proof is an immediate implication of Theorem 3.2.1.5 and the Theorems 3.2.3.1 and 3.2.3.2. □

3.3 Calculating the joint densities

3.3.1 Abstract formulae for the densities

Let us go back to the notations of Chapter 2 and let X denote a diffusion process defined by the stochastic differential equation (2.1). For the upcoming discussion, we implicitly assume that the coefficients μ and σ of X are sufficiently smooth. Let $\mathcal{D} \subset \mathbb{R}$ be a bounded open interval and let $t > 0$. From the definition of the transition probability density $p^{\mathcal{D}}$ of the killed diffusion $X^{\mathcal{D}}$, it follows that, for any Borel set $B \in \mathcal{B}(\mathcal{D})$, we have

$$\mathbb{P}_x[X_t \in B,\ t < \zeta_{\mathcal{D}}] = \int_B p^{\mathcal{D}}(t, x, y)dy, \tag{3.52}$$

where ζ_D denotes the first exit time from $\mathcal{D}$, defined by (2.13). The latter probability can be expressed in a different way. Obviously,

$$\mathbb{P}_x[X_t \in B,\ t < \zeta_{\mathcal{D}}] = \mathbb{P}_x[X_s \in \mathcal{D}, \forall s \le t;\ X_t \in B]. \tag{3.53}$$

This is an immediate consequence of the fact that at time t there are two possibilities: either the process X starting at $x \in \mathcal{D}$ crossed the boundary $\partial\mathcal{D}$ and was killed or it is still in the interior of $\mathcal{D}$.

Let $l, h \in \mathbb{R}$ with $l < h$ and set $\mathcal{D} = (l, h)$. For $t > 0$, we write $L_t = \inf_{0 \le s \le t} X_s$ and $H_t = \sup_{0 \le s \le t} X_s$. Equation (3.53) can be used to derive the joint probability of (H_t, L_t, X_t). According to (3.52) and (3.53), for $B \in \mathcal{B}((l, h))$, we find

$$\mathbb{P}_x\Big[X_s \in (l,h) \text{ for } s \le t,\ X_t \in B\Big] = \mathbb{P}_x\left[\ l < \inf_{0 \le s \le t} X_s \le \sup_{0 \le s \le t} X_s < h;\ X_t \in B\right]. \tag{3.54}$$

By the preceding deliberations and by the results of Chapter 2 we have found a powerful means to characterize the quantity (3.54) for a class of diffusion processes.

Now, let us assume that the joint densities of both (H_t, L_t) and (H_t, L_t, X_t) exist, that is we assume that the laws of the respective vectors are absolutely continuous with respect to the Lebesgue measure. Conditions for this to hold have been stated in the previous Section 3.2. We denote the joint densities with $f_{(H,L)}(t, x, h, l)$ and $f_{(H,L,X)}(t, x, h, l, y)$, respectively. From equation (3.54) it follows directly that $f_{(H,L)}$ and $f_{(H,L,X)}$ must satisfy

$$f_{(H,L)}(t, x, h, l) = -\frac{\partial^2}{\partial h \partial l} \int_{\mathcal{D}} p^{\mathcal{D}}(t, x, y) dy \tag{3.55}$$

and

$$f_{(H,L,X)}(t, x, h, l, y) = -\frac{\partial^2}{\partial h \partial l} p^{\mathcal{D}}(t, x, y), \tag{3.56}$$

respectively. Here, we use the notation $\partial^2 / \partial h \partial l$ to denote the weak derivatives with respect to l and h. Let us state these results in the following two propositions.

Proposition 3.3.1.1. *Let us consider a diffusion X defined by the stochastic differential equation* (2.1) *and let X start in $X_0 = x$. Suppose that the coefficients μ and σ of X satisfy the assumptions of Theorem 3.2.3.4 in the previous Section 3.2. Then, for $t > 0$, the joint density $f_{(H,L,X)}(t, x, h, l, y)$ of (H_t, L_t, X_t) exists, and it satisfies*

$$f_{(H,L,X)}(t, x, h, l, y) = -\frac{\partial^2}{\partial h \partial l} p^{\mathcal{D}}(t, x, y). \tag{3.57}$$

Here, $\mathcal{D} = (l, h)$ and $p^{\mathcal{D}}$ denotes the transition probability density of the associated killed process $X^{\mathcal{D}}$.

Proof. The assumptions of Theorem 3.2.3.4 guarantee the existence of a transition density $p^{\mathcal{D}}(t, x, y)$ for the process X killed at the boundary of $\mathcal{D} = (l, h)$. Also see the discussion at the beginning of Section 2.3. According to our above deliberations, for any

Borel set $B \in \mathcal{B}(\mathcal{D})$, we have

$$\mathbb{P}_x[X_t \in B,\ \zeta_D > t] = \int_B p^{\mathcal{D}}(t,x,y)dy. \tag{3.58}$$

The left hand side of (3.58) coincides with the right hand side of (3.54). Due to our assumptions, the joint density of (H_t, L_t, X_t), conditional on $X_0 = x$, exists and hence we are allowed to differentiate the right hand side of (3.58) with respect to (h, l, y) in the weak sense. □

Proposition 3.3.1.2. *Again we suppose that the coefficients μ and σ of X satisfy the assumptions of Theorem 3.2.3.4 in the previous Section 3.2. Then, for $t > 0$, the joint density $f_{(H,L)}(t,x,h,l)$ of the vector (H_t, L_t) exists and it satisfies*

$$f_{(H,L)}(t,x,h,l) = -\frac{\partial^2}{\partial h \partial l}\int_{\mathcal{D}} p^{\mathcal{D}}(t,x,y)dy. \tag{3.59}$$

Again we set $\mathcal{D} = (l,h)$ and again $p^{\mathcal{D}}$ denotes the transition probability density of the associated killed process $X^{\mathcal{D}}$.

Proof. Setting $B = \mathcal{D}$ in formula (3.58), we get

$$\mathbb{P}_x[\zeta_{\mathcal{D}} > t] = \int_{\mathcal{D}} p^{\mathcal{D}}(t,x,y)dy. \tag{3.60}$$

The result follows by the same reasoning we conducted to prove Proposition 3.3.1.1. □

Remark 3.3.1.3. Once again, we want to emphasize an important fact. Let $\mathcal{A}$ denote the infinitesimal generator of the diffusion process X defined by the stochastic differential equation (2.1). The results of Section 2.3 show that a sufficiently smooth solution to the partial differential equation

$$\frac{\partial}{\partial t}u(t,x) = \mathcal{A}u(t,x), \quad \forall x \in \mathcal{D},\ t \geq 0, \tag{3.61}$$

with initial condition

$$u(0,x) = \delta(x-y), \tag{3.62}$$

and boundary conditions

$$u(t,x) = 0, \quad \forall x \in \partial\mathcal{D},\, t \geq 0, \tag{3.63}$$

coincides with the transition probability density $p^{\mathcal{D}}(t,x,y)$ of the diffusion X killed at the boundary of $\mathcal{D} = (l,h)$. To our knowledge, the assumptions of Theorem 3.2.3.4 are not sufficient to guarantee the smoothness of $p^{\mathcal{D}}(t,x,y)$ in (t,x). Slightly stronger assumptions have to be imposed. For example, Ladyshenskaja et al. [47] proved that there is a solution $p^{\mathcal{D}}$ to the initial/boundary value problem (3.61), (3.62), (3.63) in the classical sense if the coefficients μ and σ of X satisfy Condition 2.3.0.12 in Chapter 2.

The following corollary is an immediate consequence of Proposition 3.3.1.2.

Corollary 3.3.1.4. *Let us write $p(t,x,y;h,l)$ instead of $p^{\mathcal{D}}(t,x,y)$. On the assumptions of Theorem 3.3.1.2 we have*

$$f_{(H,L)}(t,x,h,l) = \int_l^h \frac{\partial^2}{\partial h \partial l} p(t,x,y;h,l)dy + p_2(t,x,l;h,l) - p_1(t,x,h;h,l), \tag{3.64}$$

where

$$p_1(t,x,h;h,l) = \frac{\partial}{\partial z} p(t,x,l;z,l)|_{z=h} \quad \text{and} \quad p_2(t,x,l;h,l) = \frac{\partial}{\partial z} p(t,x,h;h,z)|_{z=l}. \tag{3.65}$$

Proof. Use the chain rule. □

Let us close this section with a quite obvious observation. Recalling the definition of the first exit time $\zeta_{\mathcal{D}}$ given in (2.13), we are able to state that the joint distribution of (H,L) must satisfy

$$\mathbb{P}_x[\zeta_{\mathcal{D}} > t] = 1 - \mathbb{P}_x[\zeta_{\mathcal{D}} \leq t] = \mathbb{P}_x\Big[l < \inf_{0\leq s\leq t} X_s \leq \sup_{0\leq s\leq t} X_s < h\Big]. \tag{3.66}$$

For X starting in x, let $f_\tau(x,\cdot)$ denote the density of the stopping time of $\zeta_{\mathcal{D}} = \inf\{t > 0 | X_t \notin (l,h)\}$. From our previous results we can easily infer the following corollary.

Corollary 3.3.1.5. *Let the assumptions of Proposition 3.3.1.2 be satisfied. Moreover, assume that the stopping time ζ_D has a density f_τ with respect to the Lebesgue measure. Then the joint density $f_{(H,L)}$ of the processes H and L at time t satisfies*

$$f_{(H,L)}(t,x,h,l) = \frac{\partial^2}{\partial h \partial l} \int_0^t f_\tau(x,s)\, ds. \tag{3.67}$$

Proof. Since

$$\mathbb{P}_x[\zeta_{\mathcal{D}} \leq t] = \int_0^t f_\tau(x,s)\, ds, \tag{3.68}$$

the result follows immediately. □

Concluding remarks Let us briefly sum up the results we have found so far. We considered a one-dimensional, time-homogeneous diffusion process X on $\mathbb{R}$, given by a stochastic differential equation of the form (2.1). For this process we proved an existence result for the joint densities $f_{(H,L)}$ and $f_{(H,L,X)}$ of the random vectors (H,L) and (H,L,X). The only condition necessary to impose upon the drift coefficient μ and the diffusion coefficient σ was that both coefficients should be differentiable with uniformly bounded first derivative and that σ should be uniformly elliptic.

The statements of Proposition 3.3.1.1 and Proposition 3.3.1.2 tell us how to calculate the densities - at least theoretically. Explicit solutions to the initial/boundary value problem (3.61)-(3.63), that characterizes the transition density $p^{\mathcal{D}}$ of the killed diffusion $X^{\mathcal{D}}$, do not exist in general. Thus, either extensive simulations of the triplet (H, L, X) or numerical methods for partial differential equations are necessary to find approximations to the density functions $f_{(H,L)}$ and $f_{(H,L,X)}$.

3.3.2 Calculating densities using Laplace transforms.

Usually, an explicit representation of the joint density (3.59) does not exist. Only if X is a Brownian motion with drift, or if X is an Ornstein-Uhlenbeck process, explicit formulae are known. But even for these simple processes the calculation of the densities is very cumbersome, as we will see in the examples given below. In this section we propose a relatively easy way to calculate the Laplace transform of the density (3.59). Recall the definition of the stopping time $\zeta_{\mathcal{D}}$ given in (2.13). Here, $\mathcal{D}$ is assumed to be an interval $\mathcal{D} = (l, h)$ with $l, h \in \mathbb{R},\ l < h$. For a diffusion X starting in $x \in \mathcal{D}$, we denote the density of $\zeta_{\mathcal{D}}$ with $f_\tau(x, \cdot)$. It turns out that the Laplace transform of the density $f_{(H,L)}$ can be expressed in terms of the Laplace transform of the density f_τ. In general, the Laplace transform will have a much simpler form than the density (3.59) itself. The results presented in this section rely on basic properties initially found by Darling and Siegert [18]. First, let us consider the following condition.

Condition 3.3.2.1. *Let X denote the process* (2.1)*. Assume that the transition density p of X exists and that, for all $y \in \mathbb{R}$, it satisfies the equation*

$$\frac{\partial}{\partial t} p(t, x, y) = \mu(x) \frac{\partial}{\partial x} p(t, x, y) + \frac{1}{2} \sigma(x)^2 \frac{\partial^2}{\partial x^2} p(t, x, y), \quad \forall x \in \mathbb{R},\ t \geq 0, \tag{3.69}$$

with initial condition $p(0, x, y) = \delta(x - y)$. Here, $\delta(\cdot - y)$ denotes the Dirac measure of y.

The crucial point is that the following theorem holds.

Theorem 3.3.2.2. *Let $l, h \in \mathbb{R}$, $l < h$. Assume that Condition 3.3.2.1 above is satisfied. For the Laplace transform $\hat{f}_\tau$ of the density of the stopping time $\zeta_{(l,h)} = \tau_{(l,h)^c} = \inf\{t > 0 | X_t \notin (l, h)\}$ we have*

$$\mathbb{E}_x \left[e^{-\xi \tau} \right] = \hat{f}_\tau(x, \xi) = \frac{v(x)(u(h) - u(l)) - u(x)(v(h) - v(l))}{u(h)v(l) - u(l)v(h)}, \tag{3.70}$$

where the functions u and v can be chosen as any two linearly independent solutions to the differential equation

$$\mu(x) \frac{\partial}{\partial x} w(x) + \frac{1}{2} \sigma(x)^2 \frac{\partial^2}{\partial x^2} w(x) - \xi w(x) = 0, \quad \forall x \in \mathbb{R}. \tag{3.71}$$

Proof. We will only give a sketch of the proof. The proof is mainly based on the fact that, if p satisfies (3.69), its Laplace transform

$$\hat{p}(\xi, x, y) = \int_0^\infty e^{-\xi t} p(t, x, y) dt \tag{3.72}$$

clearly satisfies the differential equation

$$\xi \hat{p}(\xi, x, y) = \mu(x) \frac{\partial}{\partial x} \hat{p}(\xi, x, y) + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} \hat{p}(\xi, x, y), \quad \forall x \in \mathbb{R}. \tag{3.73}$$

Note that this differential equation coincides with the differential equation (3.71). Now, it is necessary to split the problem. Absorption in the upper and in the lower barrier has to be considered separately. Then a sophisticated combination of both cases yields the overall result. For a detailed proof see Darling and Siegert [18] or Chapter 3.4 in Bharucha-Reid [8]. □

We close this section with an application of the previous theorem. We assume that $\mathcal{D} = (l, h)$, $l < h$, is an interval. From Fubini's theorem it follows that, for any bounded function $g \in L^1$,

$$\int_0^\infty e^{-\xi t} \left(\int_0^t g(s) ds \right) dt = \frac{\hat{g}(\xi)}{\xi}, \quad \xi > 0. \tag{3.74}$$

Hence, for fixed x, the Laplace transform of the function $b_\mathcal{T}(x, t) = \mathbb{P}_x[\zeta_\mathcal{D} \leq t]$ is given by

$$\hat{b}_\mathcal{T}(x, \xi) = \frac{\hat{f}_\mathcal{T}(x, \xi)}{\xi}, \tag{3.75}$$

where $f_\mathcal{T}$ is the density of the stopping time $\zeta_\mathcal{D} = \tau_{\mathcal{D}^c} = \inf\{t > 0 | X_t \notin \mathcal{D}\}$. On the other hand, if the joint density $f_{(H,L)}$ of (H_t, L_t) exists for all $t \geq 0$, then we are allowed to write

$$1 - b_\mathcal{T}(x, t) = \int_l^x \int_x^h f_{(H,L)}(t, x, a, b) da db. \tag{3.76}$$

And consequently, by Fubini's theorem, we find that, for $\xi > 0$,

$$\frac{1}{\xi} - \hat{b}_\mathcal{T}(x, \xi) = \int_l^x \int_x^h \hat{f}_{(H,L)}(\xi, x, a, b) da db. \tag{3.77}$$

Overall, we obtain the following relation

$$\hat{f}_{(H,L)}(x, \xi) = \frac{\partial^2}{\partial h \partial l} \frac{\hat{f}_\mathcal{T}(x, \xi)}{\xi}. \tag{3.78}$$

3.4 Examples

3.4.1 One-dimensional Brownian motion

Let us consider a one-dimensional Brownian motion with deterministic drift $\mu \in \mathbb{R}$ and diffusion coefficient $\sigma > 0$ that starts at $X_0 = x$. Such a process satisfies the stochastic differential equation

$$dX_t = \mu\, dt + \sigma dB_t, \quad X_0 = x,\ t \geq 0, \tag{3.79}$$

where B denotes the standard Brownian motion of $\mathbb{R}$. We want to calculate the joint probability of the maximum and the minimum of X at time t. Therefore, we have to solve Kolmogorov's forward equation with Dirichlet boundary conditions. In other words, for $l, h \in \mathbb{R}$ with $l < x < h$, we have to find a solution to

$$\frac{\partial}{\partial t} u(t,y) = -\mu \frac{\partial}{\partial y} u(t,y) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial y^2} u(t,y), \quad \forall (t,y) \in \mathbb{R}_+ \times \mathbb{R}, \tag{3.80}$$

with initial condition

$$u(0,y) = \delta(y - x), \tag{3.81}$$

and with the boundary conditions

$$u(t,h) = 0 \quad \text{and} \quad u(t,l) = 0, \quad \forall t \geq 0. \tag{3.82}$$

The partial differential equation described by (3.80), (3.81) and (3.82) was solved explicitly by Dominé [20]. Henceforth, we will denote the solution to this initial/boundary value problem by $p_{\mu,\sigma^2}^{(l,h)}(t,x,y)$.

In the case of $\mu \equiv 0$ a solution is particularly simple to obtain. It can easily be shown that for the Hilbert space $L^2([l,h])$ the system

$$u_k(x) = \sqrt{\frac{2}{h-l}} \sin\left(\pi k \frac{x-l}{h-l}\right), \quad k \in \mathbb{N}, \tag{3.83}$$

is an orthonormal basis of eigenfunctions with respect to the infinitesimal generator $\mathcal{A} = \frac{\sigma^2}{2}\frac{d^2}{dx^2}$. The corresponding eigenvalues are given by

$$\lambda_k = \frac{\sigma^2 \pi^2 k^2}{2(h-l)^2}, \quad k \in \mathbb{N}. \tag{3.84}$$

Besides, each function in this system evidently satisfies the Dirichlet boundary conditions (3.82). A separation approach shows that

$$p_{0,\sigma^2}^{(l,h)}(t,x,y) = \sum_{k=0}^{\infty} \exp(-\sigma^2 \lambda_k t) u_k(x) u_k(y)$$

$$= \sum_{k=0}^{\infty} \exp\left(-\frac{\sigma^2 t\pi^2 k^2}{2(h-l)^2}\right) \sin\left(\pi k \frac{x-l}{h-l}\right) \sin\left(\pi k \frac{y-l}{h-l}\right) \tag{3.85}$$

is a solution to the initial/boundary value problem we are interested in.

The outline at the beginning of Section 3.3.1 tells us that

$$\mathbb{P}_x[l < L_t,\, H_t < h] = \int_l^h p_{0,\sigma^2}^{(l,h)}(t,x,y)dy, \tag{3.86}$$

where we set

$$H_t = \sup_{0\le s\le t}\{\sigma B_s\} \quad \text{and} \quad L_t = \inf_{0\le s\le t}\{\sigma B_s\}. \tag{3.87}$$

By means of (3.85), we are able to derive the joint density of H_t and L_t. We denote this density with $f_{(H,L)}$. Before we calculate it, let us rewrite $p_{0,\sigma^2}^{(l,h)}$ in order to find a representation of $f_{(H,L)}$ that is more convenient. From formulae (3.83), (3.84) and (3.85) we infer that, for $x, y \in [l,h]$, we have

$$p_{0,\sigma^2}^{(l,h)}(t,x,y) = \frac{1}{h-l} p_{0,1}^{(0,1)}\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}, \frac{y-l}{h-l}\right), \tag{3.88}$$

where

$$p_{0,1}^{(0,1)}(t,x,y) = \sum_{k=1}^{\infty} 2\exp(-k^2\pi^2 t/2)\sin(k\pi x)\sin(k\pi y). \tag{3.89}$$

We find the following equation

$$\begin{aligned} \int_l^h p_{0,\sigma^2}^{(l,h)}(t,x,y)dy &= \frac{1}{h-l}\int_l^h p_{0,1}^{(0,1)}\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}, \frac{y-l}{h-l}\right)dy \\ &= \int_0^1 p_{0,1}^{(0,1)}\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}, y\right)dy. \end{aligned} \tag{3.90}$$

For $\alpha,\ \beta \in \mathbb{N}$, we set

$$\frac{\partial^\alpha}{\partial t^\alpha}\frac{\partial^\beta}{\partial x^\beta} p_{0,1}^{[0,1]}\left(\cdot,\cdot,y\right)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)} := \frac{\partial^\alpha}{\partial u^\alpha}\frac{\partial^\beta}{\partial z^\beta} p_{0,1}^{[0,1]}\left(u,z,y\right)\Big|_{u=\frac{\sigma^2 t}{(h-l)^2},\, z=\frac{x-l}{h-l}}. \tag{3.91}$$

By making use of this notation, and by differentiating equation (3.90) with respect to h, we obtain

$$\frac{\partial}{\partial h}\int_l^h p_{0,\sigma^2}^{(l,h)}(t,x,y)dy = \frac{-2\sigma^2 t}{(h-l)^3}\int_0^1 \frac{\partial}{\partial t} p_{0,1}^{(0,1)}\left(\cdot,\cdot,y\right)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)} dy$$

$$- \frac{x-l}{(h-l)^2} \int_0^1 \frac{\partial}{\partial y} p_{0,1}^{(0,1)}(\cdot,\cdot,y) \Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)} dy. \tag{3.92}$$

By differentiating the latter expression with respect to l, one eventually obtains the joint density $f_{(H,L)}(t,x,h,l;\sigma^2)$ of $(H_t, L_t) = \sigma\big(\sup_{0\le s\le t} B_s, \inf_{0\le s\le t} B_s\big)$, conditional on $X_0 = x$. It is given by

$$\begin{aligned}
& f_{(H,L)}(t,x,h,l;\sigma^2) \\
= \; & -\frac{\partial}{\partial l}\frac{\partial}{\partial h}\int_l^h p_{0,\sigma^2}^{(l,h)}(t,x,y)dx \\
= \; & \frac{6\sigma^2 t}{(h-l)^4}\int_0^1 \left(\frac{\partial}{\partial t} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)}\right) dy \\
+ \; & \frac{4\sigma^4 t^2}{(h-l)^6}\int_0^1 \left(\frac{\partial^2}{\partial t^2} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)}\right) dy \\
+ \; & \frac{2\sigma^2 t(x-h)}{(h-l)^5}\int_0^1 \left(\frac{\partial}{\partial t}\frac{\partial}{\partial x} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)}\right) dy \\
- \; & \frac{(h-l)^2 - 2(x-l)(h-l)}{(h-l)^4}\int_0^1 \left(\frac{\partial}{\partial x} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)}\right) dy \\
+ \; & \frac{2\sigma^2 t(x-l)}{(h-l)^5}\int_0^1 \left(\frac{\partial}{\partial t}\frac{\partial}{\partial x} p_{0,1}^{(0,1)}(\cdot,\cdot,x)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)}\right) dy \\
- \; & \frac{x-l}{(h-l)^2}\frac{x-h}{(h-l)^2}\int_0^1 \left(\frac{\partial^2}{\partial x^2} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)}\right) dy.
\end{aligned} \tag{3.93}$$

Let us mention that, for $\alpha, \beta \in \mathbb{N}_0$, we have the following two representations. If $\beta = 0$ or if β is an even number, we have

$$\begin{aligned}
& \int_0^1 \frac{\partial^\alpha}{\partial t^\alpha}\frac{\partial^\beta}{\partial x^\beta} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)} dx \\
& = \sum_{k=1}^\infty \frac{2}{2^\alpha}\frac{k^{2\alpha+\beta}\pi^{2\alpha+\beta}t^\alpha}{(h-l)^{2\alpha+\beta}} \exp\left(-k^2\pi^2\frac{\sigma^2 t}{2(h-l)^2}\right)(1+(-1)^{k+1})\sin\left(k\pi\frac{x-l}{h-l}\right).
\end{aligned} \tag{3.94}$$

And if β is an odd number, we have

$$\begin{aligned}
& \int_0^1 \frac{\partial^\alpha}{\partial t^\alpha}\frac{\partial^\beta}{\partial x^\beta} p_{0,1}^{(0,1)}(\cdot,\cdot,y)\Big|_{\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}\right)} dy \\
& = \sum_{k=1}^\infty \frac{2}{2^\alpha}\frac{k^{2\alpha+\beta}\pi^{2\alpha+\beta}t^\alpha}{(h-l)^{2\alpha+\beta}} \exp\left(-k^2\pi^2\frac{\sigma^2 t}{2(h-l)^2}\right)(1+(-1)^{k+1})\cos\left(k\pi\frac{x-l}{h-l}\right).
\end{aligned} \tag{3.95}$$

This completes the discussion of the case where $\mu \equiv 0$. If $\mu \not\equiv 0$, things are more complicated. We content ourselves with quoting the result. Let

$$\begin{aligned}\tilde{K}(n,x,h,l,\mu,\sigma) = &\left[\exp\left(\frac{\mu}{\sigma^2}(h-x)\right)(-1)^{n+1} + \exp\left(\frac{\mu}{\sigma^2}(l-x)\right)\right] \\ &\times \sin\left(n\pi\frac{x-l}{h-l}\right)\exp\left(-\frac{b}{\sigma^2}x\right),\end{aligned} \tag{3.96}$$

then one obtains

$$\int_l^h p_{\mu,\sigma^2}^{(l,h)}(t,x,y)dy = \sum_{n=1}^{\infty} \frac{2\sigma^4 n\pi \tilde{K}(n,x,h,l,\mu,\sigma)}{\mu^2(h-l)^2 + \sigma^4 n^2\pi^2} \exp\left(-\frac{n^2\pi^2\sigma^2 t}{2(h-l)^2} - \frac{\mu^2 t}{2\sigma^2}\right). \tag{3.97}$$

The joint density $f_{(H,L)}$ will not be derived for this case, since this is not very demonstrative. Instead, let us state that for X starting in $X_0 = x$ the density $f_\tau(x,t)$ of the stopping time $\tau = \inf\{t > 0 \,:\, X_t \notin (l,h)\}$ is given by

$$f_\tau(x,t) = f_\tau(x,t;h,l) = \sum_{n=1}^{\infty} \frac{\sigma^2 n\pi \tilde{K}(n,x,h,l,\mu,\sigma)}{(h-l)^2} \exp\left(-\frac{n^2\pi^2\sigma^2 t}{2(h-l)^2} - \frac{\mu^2 t}{2\sigma^2}\right). \tag{3.98}$$

This follows by direct calculations. Additionally, Dominé [20] showed that, letting $l \to -\infty$ in the latter formula, one obtains the inverse Gaussian distribution

$$\lim_{l\to-\infty} f_\tau(x,t;h,l) = \frac{h-x}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{(h-x-\mu t)^2}{2\sigma^2 t}\right). \tag{3.99}$$

Let $\Phi : \mathbb{R} \to \mathbb{R}$ denote the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{s^2}{2}\right) ds. \tag{3.100}$$

From (3.99) it can directly be derived that the process $H_t = \sup_{0\le s\le t}\{\mu t + \sigma B_s\}$ satisfies

$$\begin{aligned}&\mathbb{P}_x[H_t < h] \\ &= 1 - \Phi\left(-\frac{h-x-\mu t}{\sqrt{t}\sigma}\right) - \exp\left(\frac{2(h-x)\mu}{\sigma^2}\right)\Phi\left(-\frac{h-x+\mu t}{\sqrt{t}\sigma}\right).\end{aligned} \tag{3.101}$$

Consequently, we find

$$\begin{aligned}&\mathbb{P}_x[H_t \in dh] \\ &= \sqrt{\frac{2}{\pi t\sigma^2}} \exp\left(-\frac{(h-x-\mu t)^2}{2\sigma^2 t}\right) - \frac{2\mu}{\sigma^2}\exp\left(\frac{2\mu(h-x)}{\sigma^2}\right)\Phi\left(-\frac{h-x+\mu t}{\sqrt{t}\sigma}\right).\end{aligned} \tag{3.102}$$

Let us consider the particular case where the starting point x equals zero and where the coefficients are $\mu = 0$ and $\sigma = 1$. From the right hand side of the previous formula we infer that the process $H_t = \sup_{0\le s\le t} B_s$ has the following density

$$\mathbb{P}_0[H_t \in dh] = \sqrt{\frac{2}{\pi t}} \exp\left(-\frac{h^2}{2t}\right) dh. \tag{3.103}$$

Clearly, this density and the distribution of the running maximum could have been derived by means of the reflection principle as well, see e.g. Karatzas and Shreve [43].

We conclude this paragraph by mentioning that one would have obtained formula (3.98) by simply inverting the Laplace transform $\hat{f}_\tau$ of f_τ, which is given by

$$\hat{f}_\tau(0,\xi) = \hat{f}_\tau(0,\xi;h,l) = \frac{(e^{\zeta_1 l} - e^{\zeta_2 l}) - (e^{\zeta_1 h} - e^{\zeta_2 h})}{(e^{\zeta_2 h + \zeta_1 l} - e^{\zeta_2 l + \zeta_1 h})}, \tag{3.104}$$

with

$$\zeta_1 = \frac{-\mu + \sqrt{\mu^2 + 2\sigma^2\xi}}{\sigma^2}, \quad \zeta_2 = \frac{-\mu - \sqrt{\mu^2 + 2\sigma^2\xi}}{\sigma^2}. \tag{3.105}$$

Recall the discussions following Theorem 3.3.2.2 above or see the results of Darling and Siegert [18].

3.4.2 The one-dimensional Ornstein-Uhlenbeck process

In this section we will be concerned with the one-dimensional Ornstein-Uhlenbeck process. Such a process satisfies the stochastic differential equation

$$dX_t = -\beta X_t dt + \sigma dB_t, \quad X_0 = x, \ t \ge 0, \tag{3.106}$$

with constant coefficients $-\beta < 0$ and $\sigma > 0$. From Itô's formula it can easily be derived that the process

$$X_t = xe^{-\beta t} + \sigma \int_0^t e^{-\beta(t-s)} dB_s, \quad 0 \le t < \infty, \tag{3.107}$$

yields a solution to (3.106). Let $(l,h) \subset \mathbb{R}$ with $l < h$. The transition probability density $\widetilde{p}(t,x,y)$ of the process (3.107) killed at the boundary of (l,h) must satisfy the partial differential equation

$$\frac{\partial \widetilde{p}(t,x,y)}{\partial t} = \beta \frac{\partial}{\partial y}\Big\{ y\widetilde{p}(t,x,y)\Big\} + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial y^2}\widetilde{p}(t,x,y), \tag{3.108}$$

with initial condition

$$\widetilde{p}(t,x,y) = \delta(x-y), \tag{3.109}$$

and with boundary conditions

$$\widetilde{p}(t,x,y) = 0, \text{ for } y = l, h, \text{ and } \forall t \geq 0. \tag{3.110}$$

A solution to this differential equation was found by Sweet and Hardin [70]. We are going to reproduce this solution in the sequel.

For $l < h$, we set $x = s_x(\sigma/2\beta)^{1/2}$, $y = s_y(\sigma/2\beta)^{1/2}$ and $s_l = l(2\beta/\sigma)^{1/2}$, $s_h = h(2\beta/\sigma)^{1/2}$. Then, for $t \geq 0$,

$$\begin{aligned}
&\widetilde{p}(t, s_x, s_y) \\
&= \sum_{n=1}^{\infty} \exp(-\beta\lambda_n t)\exp(-\frac{1}{4}(s_y^2 - s_x^2))Z(\lambda_n, s_y)Z(\lambda_n, s_x)\left(\int_{s_l}^{s_h} Z^2(\lambda_n, s)ds\right)^{-1},
\end{aligned} \tag{3.111}$$

where $s_l \leq s_x$, $s_y \leq s_h$. Of course, formula (3.111) needs some further explanations. First, let

$$y_e(\lambda, s) = \exp(-s^2/4)N\left(-\frac{1}{2}\lambda, \frac{1}{2}, \frac{1}{2}s^2\right), \tag{3.112}$$

$$y_0(\lambda, s) = s\exp(-s^2/4)N\left(-\frac{1}{2}\lambda + \frac{1}{2}, \frac{3}{2}, \frac{1}{2}s^2\right). \tag{3.113}$$

The expression $N(a,b,z)$ denotes Kummer's function which is defined by

$$N(a,b,z) = 1 + \frac{az}{b} + \frac{(a)_2 z^2}{(b)_2 2!} + \ldots + \frac{(a)_n z^n}{(b)_n n!} + \ldots \tag{3.114}$$

with

$$(a)_n = a(a+1)(a+2)\cdot\ldots\cdot(a+n-1), \quad (a)_0 = 1. \tag{3.115}$$

Kummer's function (3.114) satisfies the second order ordinary differential equation,

$$z\frac{d^2}{dz^2}u + (b-z)\frac{d}{dz}u - au = 0. \tag{3.116}$$

For further details about $N(\cdot,\cdot,\cdot)$, see e.g. Abramowitz and Stegun [1]. But let us come back to the core issue. For each $n \in \mathbb{N}$, the eigenvalues λ_n in formula (3.111) satisfy the relation

$$y_e(\lambda_n, s_h)y_0(\lambda_n, s_l) + y_0(\lambda_n, s_h)y_e(\lambda_n, s_l) = 0, \tag{3.117}$$

and the function Z in (3.111) is given by

$$Z(\lambda_n, s) = y_e(\lambda_n, s) - \frac{y_e(\lambda_n, s_h)}{y_0(\lambda_n, s_h)}y_0(\lambda_n, s). \tag{3.118}$$

With this at hand, we would theoretically be able to calculate the joint density of H and L. But the calculations are quite tedious and do not lead to an instructive result. Instead we content ourselves with calculating the Laplace transform of the density $f_\mathcal{T}$ of the stopping time $\zeta_{(l,h)} = \tau_{(l,h)^c}$ for $\beta = \sigma = 1$. Whittaker and Watson [71] have proved that linearly independent solutions to

$$\frac{d^2}{dx^2}w(x) - x\frac{d}{dx}(x) = \xi w(x), \quad \forall x \in \mathbb{R}, \tag{3.119}$$

are given by

$$u(x) = e^{x^2/4}C_{-\xi}(x), \quad v(x) = e^{x^2/4}C_{-\xi}(-x). \tag{3.120}$$

Here, $C_\xi(x)$ denotes Weber's function which can also be found in Abramowitz and Stegun [1]. From (3.70) we obtain

$$\hat{f}_\mathcal{T}(x,\xi) = \hat{f}_\mathcal{T}(x,\xi;h,l) = \frac{v(x)(u(h)-u(l)) - u(x)(v(h)-v(l))}{u(h)v(l) - u(l)v(h)} \tag{3.121}$$

$$= \frac{C_{-\xi}(h) - C_{-\xi}(l) - C_{-\xi}(-h) + C_{-\xi}(-l)}{C_{-\xi}(h)C_{-\xi}(-l) - C_{-\xi}(l)C_{-\xi}(-h)}, \tag{3.122}$$

and consequently it follows that

$$\hat{f}_{(H,L)}(\xi,x,h,l) = \frac{1}{\xi}\frac{\partial^2}{\partial h \partial l}\frac{C_{-\xi}(h) - C_{-\xi}(l) - C_{-\xi}(-h) + C_{-\xi}(-l)}{C_{-\xi}(h)C_{-\xi}(-l) - C_{-\xi}(l)C_{-\xi}(-h)}, \tag{3.123}$$

see also Theorem 3.3.2.2 and the discussion afterwards.

Some more facts that are closely related to this topic are described in the article of Pedersen et al. [54]. The authors depict three different representations for the hitting time density of an Ornstein-Uhlenbeck process. And finally, let us note that the results of Chapter 7 permit to derive another, completely new, representation of the killed transition density $p^{(-\infty,h)}(t,x,y)$ for a class of diffusions that also comprises the Ornstein-Uhlenbeck model. See particularly Theorem 7.5.0.4 and Corollary 7.5.0.5.

3.4.3 Discrete approximation of diffusions

Let $(B_t,\, t \geq 0)$ denote the standard one-dimensional Brownian motion and let ξ_i be a sequence of i.i.d. random variables with $\mathbb{E}\xi_i = 0$, $\mathbb{E}\xi_i^2 = \sigma^2$. The process defined via

$$Y_t^n = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{\lfloor tn \rfloor}\xi_i + (tn - \lfloor tn \rfloor)\frac{1}{\sigma\sqrt{n}}\xi_{\lfloor tn \rfloor + 1}, \quad t \in [0,1], \tag{3.124}$$

converges to B weakly in $(\mathcal{C}([0,1],\mathbb{R}),\|\cdot\|_\infty)$ as $n\to\infty$, which we write as $Y\Rightarrow_n B$. The continuous mapping theorem tells us that

$$\sup_{0\le t\le 1} Y_t^n \Longrightarrow_n \sup_{0\le t\le 1} B_t. \tag{3.125}$$

Let $M_n=\sum_{i=1}^n \xi_i$, then obviously

$$\sup_{0\le t\le 1} Y_t^n = \frac{\sup_{0\le i\le n} M_n}{\sigma\sqrt{n}}. \tag{3.126}$$

So the limiting distribution of $\sup_{0\le i\le n} M_n/\sigma\sqrt{n}$ is the distribution of $\sup_{0\le t\le 1} B_t$. For a more detailed discussion see Billingsley [15], Section 8. The idea is to mimic this proceeding for general diffusions in lieu of Brownian motion.

Let us consider the stochastic differential equation

$$dX_t=\mu(X_t)dt+\sigma(X_t)dB_t,\quad X_0=x,\ t\ge 0. \tag{3.127}$$

We assume that the coefficients μ and σ satisfy Condition 2.3.0.12. Let $p(t,x,y)$ denote the transition probability density and let $(T_t,\ t\ge 0)$ denote the transition operators of X. Additionally, we denote the infinitesimal generator of X with $\mathcal{A}$ and its domain with $\operatorname{dom}(\mathcal{A})$. Consider the sequence of discrete time Markov processes $\{(\xi_i^n,\ i\ge 0)\}$ indexed by $n\in\mathbb{N}$ and defined by $\xi_i^n = X_{\frac{i}{n}}$. We define the approximating process Y^n via $(Y_t^n=\xi_{\lfloor tn\rfloor}^n,\ t\ge 0)$ and we will presently show that

$$Y^n \Longrightarrow_n X. \tag{3.128}$$

Here $\Rightarrow_n$ denotes weak convergence in the space of right continuous functions with left hand limits. This space is usually denoted with $D(\mathbb{R}_+,\mathbb{R})$ and it is endowed with the Skorokhod topology. Once again, for more details see e.g. Billingsley [15]. For $t\ge 0$ and $n\in\mathbb{N}$, the transition operator $T_{n,t}$ of the process Y satisfies

$$T_{n,t}f(x)=\int_{\mathbb{R}} f(y)p\left(\frac{\lfloor tn\rfloor}{n},x,y\right)dy,\quad f\in\operatorname{dom}(\mathcal{A}). \tag{3.129}$$

On Condition 2.3.0.12 it is possible to show that, for fixed $T>0$, there exist positive constants λ_1 and λ_2, depending on μ and σ only, such that

$$p(t,x,y)\le \frac{\lambda_1}{\sqrt{t}}\exp\left(-\frac{|y-x|^2}{\lambda_2 t}\right),\quad \forall (t,x,y)\in[0,T]\times\mathbb{R}^2, \tag{3.130}$$

see e.g. Friedman [28]. By dominated convergence we find that $\lim_{n\to\infty} T_{n,t}=T_t$ strongly for any $t>0$. And hence (3.128) follows directly from Theorem 19.25 in Kallenberg [41].

Since for fixed $t > 0$ the function $\sup_t : D([0,t], \mathbb{R}^d) \to \mathbb{R}$ is continuous, we are able to state an asymptotic behavior in the following sense

$$\sup_{0 \le s \le t} Y_s^n = \max_{0 \le i \le \lfloor tn \rfloor} \xi_i^n \Longrightarrow_n \sup_{0 \le s \le t} X_t. \tag{3.131}$$

More precisely, the latter limit theorem is a direct consequence of the following proposition.

Proposition 3.4.3.1. *If f is a continuous function, a sequence (f_n) converges to f for the Skorokhod topology if and only if it converges to f locally uniformly.*

Proof. See Proposition VI 1.17 on page 328 in the book of Jacod and Shiryaev [40]. □

4 Martingale Estimating Functions

4.1 Introduction & Motivation

We give a brief introduction to martingale estimating functions for discretely observed diffusion processes. For an introduction to diffusions, we make reference to Chapter 2.

We consider a diffusion model, given by the parameterized stochastic differential equation

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = x, \ t \geq 0, \tag{4.1}$$

where B is Brownian motion and μ and σ are sufficiently smooth coefficients. For simplicity's sake we assume that the process X is one-dimensional and that the parameter θ varies in a subset Θ of $\mathbb{R}$. For $t > s$ it is assumed that X_t given $X_s = x$ has a density with respect to the Lebesgue measure, which we denote by

$$y \longmapsto p(t-s, x, y; \theta), \quad y \in \mathbb{R}. \tag{4.2}$$

For every parameter $\theta \in \Theta$, we associate a solution to (4.1) with its Markov measure on the coordinate variable space. This measure will be denoted with $\mathbb{P}_{x,\theta}$. More generally, we write $\mathbb{P}_{\nu,\theta}$ if X_0 has distribution ν. The corresponding expectation operators are denoted with $\mathbb{E}_{x,\theta}$ and $\mathbb{E}_{\nu,\theta}$, respectively. Whenever there is no ambiguity, we will simply write $\mathbb{P}_\theta$ and $\mathbb{E}_\theta$. Finally we do not allow misspecified models in our analysis. This means we assume that the correct parameter θ_0 belongs to the interior of the set Θ.

Suppose that, for $0 = t_0 < t_1 < \ldots < t_n$, we are given the observations $X_{t_0}, X_{t_1}, \ldots, X_{t_n}$. The diffusion process X is a Markov process, so the likelihood function, conditional on X_0, is

$$L_n(\theta) = \prod_{i=i}^{n} p(t_i - t_{i-1}, X_{t_{i-1}}, X_{t_i}; \theta). \tag{4.3}$$

The partial derivative of $\log L_n(\theta)$ with respect to θ is called score function. It is given by

$$U_n(\theta) = \sum_{i=1}^{n} \partial_\theta \log p(t_i - t_{i-1}, X_{t_{i-1}}, X_{t_i}; \theta). \tag{4.4}$$

The maximum likelihood estimator solves the equation $U_n(\theta) = 0$. The score function $U_n(\theta)$ is a *martingale estimating function*, which means that $U_n(\theta)$ is a martingale with respect to the filtration $\mathcal{F}_n = \sigma(X_{t_i},\, i = 0, \ldots, n)$ and $\mathbb{E}_\theta[U_n(\theta)] = 0$. This can easily be seen, provided that the following interchange of differentiation and integration is allowed:

$$\begin{aligned}
&\mathbb{E}_\theta\Big[\partial_\theta \log p(t_i - t_{i-1}, X_{t_{i-1}}, X_{t_i}; \theta) \mid X_{t_1}, \ldots, X_{t_{i-1}}\Big] \\
&\qquad = \mathbb{E}_\theta\left[\frac{\partial_\theta p(t_i - t_{i-1}, X_{t_{i-1}}, X_{t_i}; \theta)}{p(t_i - t_{i-1}, X_{t_{i-1}}, X_{t_i}; \theta)} \;\Big|\; X_{t_{i-1}}\right] \\
&\qquad = \int \frac{\partial_\theta p(t_i - t_{i-1}, X_{t_{i-1}}, y; \theta)}{p(t_i - t_{i-1}, X_{t_{i-1}}, y; \theta)} p(t_i - t_{i-1}, X_{t_{i-1}}, y; \theta) dy \\
&\qquad = \partial_\theta \int p(t_i - t_{i-1}, X_{t_{i-1}}, y; \theta) dy = \partial_\theta 1 = 0.
\end{aligned} \tag{4.5}$$

In general, the density p is not known explicitly. Since tedious simulations are necessary to approximate p, there is an incentive to look for an alternative inference method. It is somewhat natural to approximate the score function by a simpler martingale estimating function. Let us consider estimating functions of the following type:

$$G_n(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{t_{i-1}}, X_{t_i}; \theta), \tag{4.6}$$

where $\Delta_i = t_i - t_{i-1}$ and the function g takes the form

$$g(\Delta, x, y; \theta) = \sum_{j=1}^{N} a_j(\Delta, x; \theta) k_j(\Delta, x, y; \theta). \tag{4.7}$$

The $k_j(\Delta, x, y; \theta),\ j = 1, \ldots, N$, are given real valued functions that, in order to make the martingale property hold, must satisfy

$$\int k_j(\Delta, x, y; \theta) p(\Delta, x, y; \theta) dy = 0, \tag{4.8}$$

for all $\Delta > 0,\ x \in \mathbb{R}$ and $\theta \in \Theta$. The real valued functions $a_j(\Delta, x; \theta),\ j = 1, \ldots, N$, are sometimes called weight functions.

We have cause to believe that a solution $\hat{\theta}$ to the equation $G_n(\theta) = 0$ is a reasonable estimator. Several authors have analyzed martingale estimating functions of the above type. On some additional assumptions about the process X, consistency and asymptotic normality of $\hat{\theta}$ can be proved. See e.g. the work of Bibby and Sørensen, [11] and [12], Kessler and Sørensen [46], Kessler [44] and [45], and Pedersen [53]. An overview of the asymptotical methods can be found in the work of Sørensen, see [66]. A more recent paper that summarizes the existing results was written by Bibby, Jacobsen and Sørensen [9].

We end this introduction by quoting two very elementary - yet crucial - examples. The first is the so-called linear estimating function. For $N = 1$, it is given by (4.7) with

$$k_1(\Delta, x, y; \theta) = y - F(\Delta, x; \theta), \tag{4.9}$$

where

$$F(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[X_\Delta] = \int y p(\Delta, x, y; \theta) dy. \tag{4.10}$$

This case was studied by Bibby and Sørensen [11]. They inferred this type of linear estimating functions from an approximation to the continuous time likelihood function. However, if $\mu(\cdot)$ is independent of θ, then $F(\Delta, x; \theta)$ is only weakly - if at all - dependent on θ. In this situation we cannot use linear martingale estimating functions and it is reasonable to consider quadratic estimating functions instead. As a result of this, we come to our second example. For $N = 2$ quadratic martingale estimating functions are given by (4.7) with

$$\begin{aligned} k_1(\Delta, x, y; \theta) &= y - F(\Delta, x; \theta), \\ k_2(\Delta, x, y; \theta) &= [y - F(\Delta, x; \theta)]^2 - \phi(\Delta, x; \theta), \end{aligned} \tag{4.11}$$

where

$$\phi(\Delta, x; \theta) = \mathrm{Var}_{x,\theta}[X_\Delta] = \int [y - F(\Delta, x; \theta)]^2 p(\Delta, x, y; \theta) dy. \tag{4.12}$$

In order to justify the choice of this particular estimating function, note that for small Δ the transition density $p(\Delta, x, y; \theta)$ is well approximated by a Gaussian density function with expectation $F(\Delta, x; \theta)$ and variance $\phi(\Delta, x; \theta)$. The transition density of the approximating Gaussian density is denoted with $q(t, x, y)$ and thus, for small Δ, we have

$$p(\Delta, x, y; \theta) \approx q(\Delta, x, y; \theta) = \frac{1}{\sqrt{2\pi\phi(\Delta, x; \theta)}} \exp\left(-\frac{1}{2\phi(\Delta, x; \theta)} (y - F(\Delta, x; \theta))^2 \right). \tag{4.13}$$

The Gaussian density q can be used to calculate an approximate likelihood function. The corresponding approximate score function is given by

$$\begin{aligned} \sum_{i=1}^{n} \Bigg\{ & \frac{\partial_\theta F(\Delta, X_{t_{i-1}}; \theta)}{\phi(\Delta, X_{t_{i-1}}; \theta)} [X_{t_i} - F(\Delta_i, X_{t_{i-1}}; \theta)] \\ & + \frac{\partial_\theta \phi(\Delta, X_{t_{i-1}}; \theta)}{2\phi^2(\Delta, X_{t_{i-1}}; \theta)\Delta_i} [(X_{t_i} - F(\Delta_i, X_{t_{i-1}}; \theta))^2 - \phi(\Delta, X_{t_{i-1}}; \theta)] \Bigg\}. \end{aligned} \tag{4.14}$$

Obviously, the previous formula has the structure of a quadratic martingale estimating function.

In the sequel we will replace the single observation X_{t_i} by the triplet $(H_{t_i}, L_{t_i}, X_{t_i})$, where the $H_{t_i} = \sup_{t_{i-1} \leq s \leq t_i} X_s$ denote the suprema and the $L_{t_i} = \sup_{t_{i-1} \leq s \leq t_i} X_s$ denote the infima of X on the intervals $(t_{i-1}, t_i]$, $i = 1, \ldots, n$. For the resulting new sample

$$(H_{t_i}, L_{t_i}, X_{t_i})_{i=1,\ldots,n}, \tag{4.15}$$

we will construct martingale estimating functions, which will sometimes be called *generalized martingale estimating functions*. For the most part, we will concentrate on the case of linear and quadratic estimating functions. Consistency and asymptotic normality will be stated for the resulting estimators. Moreover, we are going to present optimality criteria that were originally found by Godambe an Heyde, see [32], and we are going to apply them to our generalized martingale estimating functions. Of course, the corresponding asymptotic and optimality properties of ordinary martingale estimating functions, that are constructed from discrete observations of X alone, can be retrieved from our generalized results.

4.2 Generalized approach to martingale estimating functions

Suppose that the sampling frequency $\Delta > 0$ is fixed and that for each time interval $(\Delta(i-1), \Delta i]$, $i \in \mathbb{N}$, we are given the observation $(H_{\Delta i}, L_{\Delta i}, X_{\Delta i})$, $i \in \mathbb{N}$, where $H_{\Delta i} = \sup_{\Delta(i-1) \le s \le \Delta i} X_s$ and $L_{\Delta i} = \inf_{\Delta(i-1) \le s \le \Delta i} X_s$. In Section 3.2, we presented an existence result for the joint density

$$(h, l, y) \longmapsto f(\Delta, x, h, l, y; \theta), \quad h, l, y \in \mathbb{R},\ l \le x, y \le h, \tag{4.16}$$

of $(H_\Delta, L_\Delta, X_\Delta)$, conditional on $X_0 = x$, with respect to the Lebesgue measure. We implicitly assume that the conditions that guarantee the existence of the joint density f are satisfied for all $\theta \in \Theta$. The likelihood function for the sample vector

$$(H_\Delta, L_\Delta, X_\Delta, H_{2\Delta}, L_{2\Delta}, X_{2\Delta}, \ldots, H_{n\Delta}, L_{n\Delta}, X_{n\Delta}) \tag{4.17}$$

takes the following form, conditional on X_0,

$$L_n^{(H,L,X)}(\theta) = \prod_{i=1}^{n} f(\Delta, X_{(i-1)\Delta}, H_{i\Delta}, L_{i\Delta}, X_{i\Delta}; \theta). \tag{4.18}$$

This follows directly from the Markov property of X and from the fact that the pairs $(H_{i\Delta}, L_{i\Delta})$, $i \in \mathbb{N}$, consist of the suprema and the infima of X on the disjoint intervals $(\Delta(i-1), \Delta i]$. From (4.18) we are able to derive the score function

$$U_n^{(H,L,X)}(\theta) = \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log f(\Delta, X_{(i-1)\Delta}, H_{i\Delta}, L_{i\Delta}, X_{i\Delta}; \theta). \tag{4.19}$$

Remark 4.2.0.2. Note that, in contrast to the case of ordinary martingale estimating functions, we encounter difficulties because the density

$$(h, l, y) \longmapsto f(\Delta, x, h, l, y; \theta) \tag{4.20}$$

vanishes for $x = h = l = y$. Thus (x, x, x) is a singular point of the logarithmic derivative

$$(h, l, y) \longmapsto \frac{\partial}{\partial\theta} \log f(\Delta, x, h, l, y; \theta). \tag{4.21}$$

But clearly, for $\Delta > 0$, we have $\mathbb{P}_{x,\theta}[\, x = H_\Delta = L_\Delta = X_\Delta \,] = 0$. Therefore

$$f(\Delta, X_{(i-1)\Delta}, H_{i\Delta}, L_{i\Delta}, X_{i\Delta}; \theta) > 0, \tag{4.22}$$

a.s. for all $i \in \mathbb{N}$ and consequently $U_n^{(H,L,X)}(\theta)$ is well-defined for all $n \in \mathbb{N}$.

The same calculation as in the introduction shows that $U_n^{(H,L,X)}(\theta)$ is a martingale with respect to $\mathcal{F}_n = \sigma(X_s,\, s \le \Delta n)$. But as for the transition density p, a lot of computational effort is necessary to simulate f. We obtain an alternative to the maximum likelihood approach by replacing $U_n^{(H,L,X)}(\theta)$ with another martingale. Particularly, let us consider martingale estimating functions of the following type:

$$G_n^{(H,L,X)}(\theta) = \sum_{i=1}^{n} g(\Delta_i, X_{(i-1)\Delta}, H_{i\Delta}, L_{i\Delta}, X_{i\Delta}; \theta), \tag{4.23}$$

where

$$g(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{N} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta). \tag{4.24}$$

The real valued functions $k_j,\ j = 1, \ldots, N$, must satisfy

$$\begin{aligned} &\int_{E(x)} k_j(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta)\, dh\, dl\, dy \\ &= \int_{-\infty}^{x} \int_{x}^{\infty} \int_{l}^{h} k_j(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta)\, dy\, dh\, dl \ = \ 0 \end{aligned} \tag{4.25}$$

in order to make $G_n^{(H,L,X)}(\theta)$ a martingale. In the previous equation, the state space of the sample points was denoted with

$$E(x) = \Big\{ (h, l, y) \in \mathbb{R}^3 \,\Big|\, l \le x \le h \text{ and } l \le y \le h \Big\}. \tag{4.26}$$

We are going to focus on two special cases, namely linear and quadratic martingale estimating functions. In our generalized context, for $N = 3$, a linear estimating function is given by (4.24) with

$$\begin{aligned} k_1(\Delta, x, h, l, y; \theta) &= h - F^H(\Delta, x; \theta), \\ k_2(\Delta, x, h, l, y; \theta) &= l - F^L(\Delta, x; \theta), \\ k_3(\Delta, x, h, l, y; \theta) &= y - F^X(\Delta, x; \theta), \end{aligned} \tag{4.27}$$

where

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[\,U\,], \quad \text{for } U \in \{X_\Delta, H_\Delta, L_\Delta\}. \tag{4.28}$$

Furthermore, we obtain a quadratic martingale estimating function for $N = 9$ with

$$\begin{aligned} k_1(\Delta, x, h, l, y; \theta) &= h - F^H(\Delta, x; \theta), \\ k_2(\Delta, x, h, l, y; \theta) &= l - F^L(\Delta, x; \theta), \\ k_3(\Delta, x, h, l, y; \theta) &= y - F^X(\Delta, x; \theta), \end{aligned} \tag{4.29}$$

and

$$\begin{aligned} k_4(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)][l - F^L(\Delta, x; \theta)] - \phi_{H,L}(\Delta, x; \theta), \\ k_5(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)][y - F^X(\Delta, x; \theta)] - \phi_{H,X}(\Delta, x; \theta), \\ k_6(\Delta, x, h, l, y; \theta) &= [l - F^L(\Delta, x; \theta)][y - F^X(\Delta, x; \theta)] - \phi_{L,X}(\Delta, x; \theta), \\ k_7(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)]^2 - \phi_{H,H}(\Delta, x; \theta), \\ k_8(\Delta, x, h, l, y; \theta) &= [l - F^L(\Delta, x; \theta)]^2 - \phi_{L,L}(\Delta, x; \theta), \\ k_9(\Delta, x, h, l, y; \theta) &= [y - F^X(\Delta, x; \theta)]^2 - \phi_{X,X}(\Delta, x; \theta), \end{aligned} \tag{4.30}$$

where $F^H(\Delta, x; \theta)$, $F^L(\Delta, x; \theta)$ and $F^X(\Delta, x; \theta)$ are defined as above and

$$\phi_{U,V}(\Delta, x; \theta) = \mathrm{Cov}_{x,\theta}[\,U, V\,], \quad \text{for } U, V \in \{X_\Delta, H_\Delta, L_\Delta\}. \tag{4.31}$$

We close this section by stating two important observations.

Remark 4.2.0.3. Obviously, the functions mentioned above are very special cases. All possible combinations of the functions $\{k_1, k_2, k_3\}$ or $\{k_1, \ldots, k_9\}$ can be used to construct a linear or a quadratic estimating function, respectively. The analytical tools we present in the sequel remain essentially the same.

Remark 4.2.0.4. Our generalized martingale estimating functions are particularly interesting if we want to estimate a parameter θ in the diffusion coefficient $\sigma(\cdot, \theta)$. Various moment estimators constructed from the triplet $(H_\Delta, L_\Delta, X_\Delta)$ exist that work well in a Brownian model: let $X = \sigma B$, where B denotes the standard Brownian motion of $\mathbb{R}$, and set $H_\Delta^B = \sup_{0 \le s \le \Delta} X_s$, $L_\Delta^B = \inf_{0 \le s \le \Delta} X_s$. Possible estimators for σ or σ^2 in this model are

$$\hat{\sigma}_{ub,1} = \sqrt{\pi}(H_\Delta^B - L_\Delta^B)/2\sqrt{2\Delta}, \tag{4.32}$$

$$\hat{\sigma}_{ub,2}^2 = (H_\Delta^B - L_\Delta^B)^2/\,\Delta \log 16, \tag{4.33}$$

$$\hat{\sigma}_{RS}^2 = \Big(H_\Delta^B(H_\Delta^B - X_\Delta) + L_\Delta^B(L_\Delta^B - X_\Delta)\Big)/\Delta \tag{4.34}$$

and

$$\hat{\sigma}^2_{GK} = \Big\{0.511(H^B_\Delta - L^B_\Delta)^2 - 0.019\Big(X_\Delta(H^B_\Delta + L^B_\Delta)^2 - 2H^B_\Delta L^B_\Delta\Big) - 0.383X^2_\Delta\Big\}/\Delta\,. \tag{4.35}$$

The estimators (4.32) and (4.33) are unbiased estimators. Note that $\mathbb{E}[H^B_\Delta L^B_\Delta] = \sigma^2\Delta(1 - 2\log 2)$. This fact was proved by Rogers and Shepp [62]. The estimator (4.34) is unbiased as well. It was found by Rogers and Satchell [61]. Finally, (4.35) describes the so-called *Garman-Klass estimator*. It is the estimator with minimal variance in the class of unbiased quadratic estimators, see [30].

4.2.1 Consistency and asymptotic normality

In the present section we prove a consistency result for estimators obtained from generalized martingale estimating functions for ergodic diffusions. Moreover, we establish a result about asymptotic normality.

Let $s(x,\theta)$ denote the density of the scale measure of X:

$$s(x;\theta) = \exp\left(-2\int_{x^*}^{x} \frac{\mu(y;\theta)}{\sigma^2(y;\theta)}dy\right), \tag{4.36}$$

where $x^* \in \mathbb{R}$ is an arbitrary point. Consider the following condition, which gives a sufficient criterion for ergodicity.

Condition 4.2.1.1. *The two following properties hold for all $\theta \in \Theta$:*

$$\int_{x^*}^{\infty} s(x;\theta)dx = \int_{-\infty}^{x^*} s(x;\theta)dx = \infty \tag{4.37}$$

and

$$\int_{-\infty}^{\infty} [s(x;\theta)\sigma^2(x;\theta)]^{-1}dx = A(\theta) < \infty. \tag{4.38}$$

As we mentioned, Condition 4.2.1.1 implies that the process X is ergodic with an invariant probability measure ν_θ which has density $[A(\theta)s(x;\theta)\sigma^2(x;\theta)]^{-1}$ with respect to the Lebesgue measure. An excellent summary of results concerning ergodicity is given in Genon-Catalot et al. [31].

We define a probability measure Q^Δ_θ on $\mathbb{R}^4$ by

$$Q^\Delta_\theta(x,h,l,y) = \nu_\theta(dx) \times f(\Delta,x,h,l,y;\theta), \tag{4.39}$$

where $f(\Delta,x,h,l,y;\theta)$ is the joint density of $(H_\Delta, L_\Delta, X_\Delta)$, conditional on $X_0 = x$. Furthermore, for a function $\gamma : \mathbb{R}^4 \to \mathbb{R}$, we use the notation $Q^\Delta_\theta(\gamma) = \int \gamma dQ^\Delta_\theta$ to describe the integral with respect to Q^Δ_θ.

For a measurable function $\gamma : \mathbb{R}^4 \to \mathbb{R}$ and for fixed $\Delta > 0$ we consider the probability measure $\mathbb{Q}_\gamma = \mathbb{Q}_{\gamma,\Delta,\theta}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$\mathbb{Q}_\gamma(A) = Q_\theta^\Delta[\gamma^{-1}(A)], \quad A \in \mathcal{B}(\mathbb{R}). \tag{4.40}$$

Moreover, we define the random variables

$$\gamma_i = \gamma(X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i}), \quad i \in \mathbb{N}. \tag{4.41}$$

If X_0 has distribution ν_θ, the stochastic process $\left(\gamma_i\right)_{i\in\mathbb{N}}$ is clearly stationary. Since the sampling frequency Δ is constant, this follows from the fact that, on the one hand, the observations $X_{\Delta i}$ are equidistant and that, on the other hand, the subsequent intervals $((i-1)\Delta, i\Delta]$, from which the suprema $H_{\Delta i}$ and the infima $L_{\Delta i}$ are taken, are disjoint intervals of equal length. The law of the process $\left(\gamma_i\right)_{i\in\mathbb{N}}$ on $\mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}$ is denoted with $\mathbb{Q}_{(\gamma_i)}$. Formally, it is given by

$$\mathbb{Q}_{(\gamma_i)}(A) = Q_\theta^\Delta[(\gamma_i)_{i\in\mathbb{N}} \in A], \; A \in \mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}. \tag{4.42}$$

For $x = (x_0, x_1, x_2, \ldots) \in \mathbb{R}^\mathbb{N}$, we define the one step shift-operator T by

$$T((x_0, x_1, x_2, \ldots)) = (x_1, x_2, \ldots), \tag{4.43}$$

and for a function $g : \mathbb{R}_+ \to \mathbb{R}$ and $t > 0$ we define the operator θ_t by $\theta_t g(\cdot) = g(t + \cdot)$. The following lemma states the ergodicity of the process $\left(\gamma_i\right)_{n\in\mathbb{N}}$.

Lemma 4.2.1.2. *Let the diffusion process X be ergodic and let the function $\gamma : \mathbb{R}^4 \to \mathbb{R}$ be measurable from $\mathcal{B}(\mathbb{R}^4)$ to $\mathcal{B}(\mathbb{R})$. Let $\mathbb{Q}_{(\gamma_i)}$ be the measure given in* (4.42) *and let T be the one step shift-operator defined in* (4.43)*. Then T is ergodic on the space $\left(\mathbb{R}^\mathbb{N}, \mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}, \mathbb{Q}_{(\gamma_i)}\right)$.*

Proof. First let $\mathcal{I}$ denote the sub-σ-field of $\mathcal{F}$ defined by

$$\mathcal{I} = \{\{X \in B\} \,|\, B \in \mathcal{F};\, B = \theta_t^{-1}(B), \forall t\}. \tag{4.44}$$

Since X is ergodic, $\mathbb{P}_{\nu,\theta}[A] = 0$ or 1, for all $A \in \mathcal{I}$.

The projections π_0 and π_Δ are measurable with respect to $\mathcal{F}$. Moreover the functions

$$X \longmapsto \sup_{0\le s\le\Delta} X_s \quad \text{and} \quad X \longmapsto \inf_{0\le s\le\Delta} X_s \tag{4.45}$$

are continuous from $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ to $\mathbb{R}$. Hence, they are measurable. As a consequence, the mapping $\pi_{H,L} : \mathcal{C}(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}^4$, given by

$$X \longmapsto \pi_{H,L}(X) = (X_0, H_\Delta, L_\Delta, X_\Delta), \tag{4.46}$$

is measurable from the σ-field $\mathcal{F}$ to $\mathcal{B}(\mathbb{R}^4)$.

Obviously, we can write the sequence $(\gamma_i)_{i\in\mathbb{N}}$ as

$$(\gamma_i)_{i\in\mathbb{N}} = \big((\gamma \circ \pi_{H,L})(X), (\gamma \circ \pi_{H,L} \circ \theta_{\Delta 1})(X), (\gamma \circ \pi_{H,L} \circ \theta_{\Delta 2})(X),\ \ldots\ \big). \tag{4.47}$$

Thus, if we define the mapping $m : \mathcal{C}(\mathbb{R}_+, \mathbb{R}) \longrightarrow \mathbb{R}^{\mathbb{N}}$ by

$$X \longmapsto \big((\gamma \circ \pi_{H,L} \circ \theta_{\Delta i})(X)\big)_{i\in\mathbb{N}}, \tag{4.48}$$

then for $A \in \mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}$ we have $m^{-1}(A) \in \mathcal{F}$.

From the fact that

$$\begin{aligned} T\big((\gamma_i)_{i\in\mathbb{N}}\big) &= T\Big(\big((\gamma \circ \pi_{H,L} \circ \theta_{\Delta i})(X)\big)_{i\in\mathbb{N}}\Big) \\ &= \big((\gamma \circ \pi_{H,L} \circ \theta_{\Delta i})(\theta_\Delta X)\big)_{i\in\mathbb{N}} \\ &= \big((\gamma \circ \pi_{H,L} \circ \theta_{\Delta(i+1)})(X)\big)_{i\in\mathbb{N}}, \end{aligned} \tag{4.49}$$

it follows that, if $T^{-1}(A) = A$, the set $B = m^{-1}(A)$ must satisfy $\theta_\Delta^{-1}(B) = B$. Thus $B \in \mathcal{I}$, and hence we have

$$\mathbb{Q}_{(\gamma_i)}(A) = \mathbb{P}_{\nu,\theta}[B] = 0 \text{ or } 1. \tag{4.50}$$

The latter fact follows, since the process X is ergodic. $\square$

The following theorem turns out to be crucial.

Theorem 4.2.1.3. *Suppose Condition 4.2.1.1 holds, and let the function $\gamma : \mathbb{R}^4 \to \mathbb{R}$ satisfy $Q_\theta^\Delta(\gamma^2) < \infty$. Then*

$$\frac{1}{n}\sum_{i=1}^{n} \gamma(X_{(i-1)\Delta}, H_{i\Delta}, L_{i\Delta}, X_{i\Delta}) \longrightarrow Q_\theta^\Delta(\gamma), \tag{4.51}$$

in $L^2(\mathbb{P}_\theta)$ as $n \to \infty$. Moreover, suppose that the function

$$x \longmapsto \int_{E(x)} \gamma(x,h,l,y) f(\Delta, x, h, l, y; \theta)\, dy dh dl, \quad x \in \mathbb{R}, \tag{4.52}$$

is identically equal to zero. Then

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \gamma(X_{(i-1)\Delta}, H_{i\Delta}, L_{i\Delta}, X_{i\Delta}) \longrightarrow N\left(0, Q_\theta^\Delta(\gamma^2)\right), \tag{4.53}$$

weakly as $n \to \infty$.

Proof. Let $\pi_0 : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ be the projection $\pi_0(x_0, x_1, x_2, \ldots) = x_0$. With the notations of Lemma 4.2.1.2, we have $\int \pi_0 d\mathbb{Q}_\gamma^{\otimes\mathbb{N}} = Q_\theta^\Delta(\gamma)$. Thus (4.51) follows directly from Birkhoff's ergodic theorem, which, for example, can be found in the book of Kallenberg [41].

Condition (4.52) ensures that $(\gamma_i)_{i\in\mathbb{N}}$ is a martingale with respect to the filtration $(\mathcal{F}_i)$. Consequently, assertion (4.53) follows from the Lindeberg-Lévy theorem for martingales stated by Billingsley [14]. $\square$

Let $\theta_0 \in \Theta$ denote the true parameter. Henceforth, we will assume that the following condition is satisfied.

Condition 4.2.1.4. *There exists a parameter value $\bar{\theta}$ such that*

$$\int_{E(x)} g(\Delta, x, h, l, y; \bar{\theta}) f(\Delta, x, h, l, y; \theta_0) dhdldy = 0, \tag{4.54}$$

for all $x \in \mathbb{R}$.

This condition states that there is a parameter $\bar{\theta}$ such that $G_n(\bar{\theta})$ is a martingale under the true probability measure $\mathbb{P}_{\theta_0}$. This condition replaces the assumption that $G_n(\theta)$ is a martingale under every $\theta \in \Theta$. We have to formulate another set of conditions.

Condition 4.2.1.5.

1. *The function g is twice continuously differentiable with respect to θ for all x, h, l, y.*
2. *The functions*

$$(x, h, l, y) \longmapsto g(\Delta, x, h, l, y; \theta), \tag{4.55}$$
$$(x, h, l, y) \longmapsto \partial_\theta g(\Delta, x, h, l, y; \theta), \tag{4.56}$$
$$(x, h, l, y) \longmapsto \partial_\theta^2 g(\Delta, x, h, l, y; \theta) \tag{4.57}$$

   are locally dominated integrable with respect to the true measure $Q_{\theta_0}^{\Delta}$. Moreover, the function $(x, h, l, y) \mapsto g(\Delta, x, h, l, y; \theta)$ is in $L^2(Q_{\theta_0}^{\Delta})$ for all $\theta \in \Theta$.

3. *$\xi(\bar{\theta}) = Q_{\theta_0}^{\Delta}(\partial_\theta g(\Delta, \cdot; \bar{\theta})) \neq 0$.*

Alternatively, one can consider the following set of conditions that only includes the first derivative with respect to the parameter θ, but requires stronger assumptions about the mean value of these derivatives.

Condition 4.2.1.6.

1. *The function g is continuously differentiable with respect to θ for all x, h, l, y.*
2. *The function $(x, h, l, y) \mapsto \partial_\theta g(\Delta, x, h, l, y; \theta)$ is locally dominated integrable with respect to the true measure $Q_{\theta_0}^{\Delta}$.*
3. *$\xi(\bar{\theta}) = Q_{\theta_0}^{\Delta}(\partial_\theta g(\Delta, \cdot; \bar{\theta})) > 0$.*

Note that both Condition 4.2.1.5 and Condition 4.2.1.6 are appropriate for our purposes. It is not important which one we consider. Basically they are equivalent. Furthermore, note that for the case of ordinary martingale estimating functions similar sets of conditions have to be imposed. For more details see e.g. Sørensen [66], Condition 3.4 and Condition 3.5. The next theorem is the main result of this section.

Theorem 4.2.1.7. *Suppose that $\bar{\theta} \in \Theta^{\circ}$ and that Condition 4.2.1.4 and either Condition 4.2.1.5 or Condition 4.2.1.6 hold. Let $G_n^{(H,L,X)}(\theta)$ be defined by* (4.23)*. An estimator $\hat{\theta}_n$ that solves the equation*

$$G_n^{(H,L,X)}(\hat{\theta}_n) = 0 \tag{4.58}$$

exists with a probability tending to one, as $n \to \infty$ and under $\mathbb{P}_{\bar{\theta}}$. Moreover,

$$\hat{\theta}_n \longrightarrow \bar{\theta}, \tag{4.59}$$

in probability under $\mathbb{P}_{\bar{\theta}}$ as $n \to \infty$ and

$$\sqrt{n}(\hat{\theta}_n - \bar{\theta}) \longrightarrow N\left(0, \frac{v(\bar{\theta})}{\xi^2(\bar{\theta})}\right), \tag{4.60}$$

weakly as $n \to \infty$, where $v(\bar{\theta}) = Q_{\theta_0}^{\Delta}(g(\Delta, \cdot; \bar{\theta})^2)$ and $\xi(\bar{\theta}) = Q_{\theta_0}^{\Delta}(\partial_\theta g(\Delta, \cdot; \bar{\theta}))$.

Remark 4.2.1.8. If $\mathbb{P}_{\bar{\theta}} = \mathbb{P}_{\theta_0}$, then $G_n^{(H,L,X)}(\theta)$ is an unbiased martingale estimating function, and $\hat{\theta}_n$ converges in probability to the true parameter value as $n \to \infty$.

Proof (of Theorem 4.2.1.7). Before we begin, note that our argumentation is similar to the one in the proof of Theorem 3.6 in Sørensen [66], which was designed for ordinary martingale estimating functions.

On Condition 4.2.1.4 it follows from Theorem 4.2.1.3 that

$$\frac{1}{\sqrt{n}} G_n^{(H,L,X)}(\bar{\theta}) \longrightarrow N(0, v(\bar{\theta})), \tag{4.61}$$

weakly as $n \to \infty$. Let

$$M_n^{(\alpha)}(\bar{\theta}) = \{\theta \in \Theta \, : \, |\theta - \bar{\theta}| \leq \alpha/\sqrt{n}\}. \tag{4.62}$$

On Condition 4.2.1.5, the theorem follows from a combination of Corollary 2.7 and Theorem 2.8 in Sørensen [66], if we can prove that for all $\alpha > 0$,

$$\sup_{\theta \in M_n^{(\alpha)}(\bar{\theta})} |G_n^{(H,L,X)}(\theta)/n| \longrightarrow 0, \tag{4.63}$$

$$\sup_{\theta \in M_n^{(\alpha)}(\bar{\theta})} |n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - \xi(\bar{\theta})| \longrightarrow 0 \tag{4.64}$$

and

$$\sup_{\theta \in M_n^{(\alpha)}(\bar{\theta})} |n^{-1}\partial_\theta^2 G_n^{(H,L,X)}(\theta) - \zeta(\bar{\theta})| \longrightarrow 0 \tag{4.65}$$

in probability as $n \to \infty$. Here, $\zeta(\bar\theta) = Q^{\Delta}_{\theta_0}(\partial^2_\theta g(\Delta,\cdot;\bar\theta))$. On Condition 4.2.1.6 one only has to show that (4.64) holds. Let us note that

$$\begin{aligned}
&\sup_{\theta \in M_n^{(\alpha)}(\bar\theta)} |n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - \xi(\bar\theta)| \\
&\qquad \le \sup_{\theta \in M_n^{(\alpha)}(\bar\theta)} |n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - \xi(\theta)| + \sup_{\theta \in M_n^{(\alpha)}(\bar\theta)} |\xi(\bar\theta) - \xi(\theta)|, \qquad (4.66)
\end{aligned}$$

where, again, $\xi(\bar\theta) = Q^{\Delta}_{\theta_0}(\partial_\theta g(\Delta,\cdot;\bar\theta))$. Each of the two terms in (4.66) tends to 0 in probability. This follows from the fact that $\xi(\theta)$ is continuous and on the other hand, for every compact subset $K \subset \Theta$,

$$\sup_{\theta\in K} |n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - \xi(\bar\theta)| \longrightarrow 0, \tag{4.67}$$

a.s. $\mathbb{P}_{\theta_0}$ as $n \to \infty$. These two facts will be proved in the sequel.

First, define the following modulus of continuity

$$k(\theta,\delta;x,h,l,y) = \sup_{|\tilde\theta-\theta|\le\delta} |\partial_\theta g(\Delta,x,h,l,y;\tilde\theta) - \partial_\theta g(\Delta,x,h,l,y;\theta)|. \tag{4.68}$$

By the dominated convergence theorem (where we use the local integrability of $\partial_\theta g$ with respect to $Q^{\Delta}_{\theta_0}$) one obtains

$$\lim_{\delta\to 0} Q^{\Delta}_{\theta_0}(k(\theta,\delta,\cdot)) = Q^{\Delta}_{\theta_0}\left(\lim_{\delta\to 0} k(\theta,\delta,\cdot)\right). \tag{4.69}$$

Now suppose that $\theta_n \to \theta$. Then

$$\begin{aligned}
|\xi(\theta_n) - \xi(\theta)| &= \left|Q^{\Delta}_{\theta_0}(\partial_\theta g(\Delta,\cdot;\theta_n)) - Q^{\Delta}_{\theta_0}(\partial_\theta g(\Delta,\cdot;\theta))\right| \\
&\le \text{const.}\, Q^{\Delta}_{\theta_0}\left(|\partial_\theta g(\Delta,\cdot;\theta_n) - \partial_\theta g(\Delta,\cdot;\theta)|\right) \\
&\le \text{const.}\, Q^{\Delta}_{\theta_0}\left(k(\theta,\delta_n,\cdot)\right) \longrightarrow 0\,, \qquad (4.70)
\end{aligned}$$

where $\delta_n = |\theta_n - \theta|$. Thus $\xi(\theta)$ is continuous.

Since $\partial_\theta g(\Delta,x,h,l,y;\bar\theta)$ is locally dominated integrable with respect to $Q^{\Delta}_{\theta_0}$, for every $\theta \in \Theta$ there is a $\delta_\theta > 0$ such that

$$k(\theta,\delta;x,h,l,y) \in L^1(Q^{\Delta}_{\theta_0}), \text{ for } 0 < \delta < \delta_\theta. \tag{4.71}$$

Let us fix $\epsilon > 0$. The function $\xi(\theta)$ is continuous. This is why, for every $\theta \in \Theta$, we can find a $\lambda_\theta \in (0,\delta_\theta]$ such that

$$|\tilde\theta - \theta| < \lambda_\theta \Longrightarrow |\xi(\tilde\theta) - \xi(\theta)| < \frac{1}{2}\epsilon \tag{4.72}$$

and

$$Q_{\theta_0}^{\Delta}(k(\theta,\lambda_\theta,\cdot)) < \frac{1}{2}\epsilon. \tag{4.73}$$

Let K be a compact subset of Θ. Then there exists a finite covering

$$K \subseteq \bigcup_{j=1}^{r} B_{\theta_j}(\lambda_{\theta_j}), \tag{4.74}$$

where $B_\theta(\lambda) = \{\tilde\theta : |\theta - \tilde\theta| < \lambda\}$ and where $\theta_1,\dots,\theta_r \in K$. For every $\theta \in K$, we can therefore choose $\theta_l,\ l \in \{1,\dots,r\}$, such that

$$|\theta - \theta_l| < \lambda_l. \tag{4.75}$$

Then, for $\theta \in K$,

$$\begin{aligned}
&\left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - \xi(\theta)\right| \\
&\quad\le \left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta_l)\right| \\
&\qquad + \left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta_l) - \xi(\theta_l)\right| + \underbrace{\left|\xi(\theta_l) - \xi(\theta)\right|}_{<\ \epsilon/2} \\
&\quad\le \frac{1}{n}\sum_{\nu=1}^{n} \Big|\partial_\theta g(\Delta, X_{\Delta(\nu-1)}, H_{\Delta\nu}, L_{\Delta\nu}, X_{\Delta(\nu)};\theta) \\
&\qquad\qquad - \partial_\theta g(\Delta, X_{\Delta(\nu-1)}, H_{\Delta\nu}, L_{\Delta\nu}, X_{\Delta(\nu)};\theta_l)\Big| \\
&\qquad + \left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta_l) - \xi(\theta_l)\right| + \frac{\epsilon}{2} \\
&\quad\le \frac{1}{n}\sum_{\nu=1}^{n} k\left(\theta_l, \lambda_l; X_{\Delta(\nu-1)}, H_{\Delta\nu}, L_{\Delta\nu}, X_{\Delta(\nu)}\right) \\
&\qquad + \left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta_l) - \xi(\theta_l)\right| + \frac{\epsilon}{2} \\
&\quad\le \left|\frac{1}{n}\sum_{\nu=1}^{n} k\left(\theta_l, \lambda_l; X_{\Delta(\nu-1)}, H_{\Delta\nu}, L_{\Delta\nu}, X_{\Delta(\nu)}\right) - Q_{\theta_0}^{\Delta}(k(\theta_l,\lambda_l,\cdot))\right| \\
&\qquad + \underbrace{\left|Q_{\theta_0}^{\Delta}(k(\theta_l,\lambda_l,\cdot))\right|}_{<\ \epsilon/2} + \left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta_l) - \xi(\theta_l)\right| + \frac{\epsilon}{2}.
\end{aligned} \tag{4.76}$$

We infer that

$$\sup_{\theta\in K}\left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta) - \xi(\theta)\right|$$

$$\leq \max_{1\leq l\leq r}\left|\frac{1}{n}\sum_{\nu=1}^{n} k\left(\theta_l,\lambda_l;X_{\Delta(\nu-1)},H_{\Delta\nu},L_{\Delta\nu},X_{\Delta(\nu)}\right)-Q_{\theta_0}^{\Delta}(k(\theta_l,\lambda_l,\cdot))\right|$$
$$+\max_{1\leq l\leq r}\left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta_l)-\xi(\theta_l)\right|+\epsilon, \tag{4.77}$$

and thus by the ergodic result stated in Theorem 4.2.1.3 we find that

$$\lim_{n\to\infty}\sup_{\theta\in K}\left|n^{-1}\partial_\theta G_n^{(H,L,X)}(\theta)-\xi(\theta)\right|\leq\epsilon, \tag{4.78}$$

almost surely for all $\epsilon > 0$. This completes the proof of the assertion. □

4.2.2 The notion of F- and A-optimality

In this section we present a notion of optimality introduced by Godambe and Heyde [32]. In the next section, the presented optimality criteria will be combined with the consistency and asymptotic results of the previous section in order to find optimal generalized linear and quadratic martingale estimating functions.

Set $\Delta = 1$ for convenience. Let us suppose that, for every $\theta \in \Theta$, the vector

$$(H_1, L_1, X_1, \ldots, H_n, L_n, X_n) \tag{4.79}$$

has a non-negative density

$$(h_1, l_1, x_1, \ldots, h_n, l_n, x_n;\theta) \longmapsto f(h_1, l_1, x_1, \ldots, h_n, l_n, x_n;\theta), \tag{4.80}$$

with respect to a dominating measure ℓ. Then the likelihood function is given by

$$L_n^{(H,L,X)}(\theta) = f(H_1, L_1, X_1, \ldots, H_n, L_n, X_n;\theta), \tag{4.81}$$

and the score function becomes

$$U_n^{(H,L,X)}(\theta) = \frac{\partial}{\partial\theta}L_n^{(H,L,X)}(\theta) = \frac{\partial_\theta f(H_1, L_1, X_1, \ldots, H_n, L_n, X_n;\theta)}{f(H_1, L_1, X_1, \ldots, H_n, L_n, X_n;\theta)}. \tag{4.82}$$

Recall Remark 4.2.0.2 about the well-definedness of the score function $U_n^{(H,L,X)}(\theta)$.

In the sequel we will drop the superscript (H, L, X) for convenience. For the rest of this chapter, we will simply write $U_n(\theta)$ and $G_n(\theta)$ to denote the generalized martingale estimating function and the generalized likelihood function, respectively. Of course the following methods hold in a much wider context. Usually they are used in situations where the likelihood function - and consequently the score function - is not known. For a given class of estimating functions $\mathcal{G}_n$, the aim is to find an element $G_n \in \mathcal{G}_n$ that is in some sense closest to the likelihood function. We consider the following condition.

Condition 4.2.2.1.

1. *$U_n(\theta) \in L^2(\mathbb{P}_\theta)$ for all $\theta \in \Theta$.*
2. *All $G_n \in \mathcal{G}_n$ satisfy $G_n \in L^2(\mathbb{P}_\theta)$ for all $\theta \in \Theta$.*

On this condition, it is reasonable to look for an element in $\mathcal{G}_n$ for which the correlation

$$\mathrm{Corr}_\theta[G_n, U_n] = \frac{\mathbb{E}_\theta[G_n(\theta)U_n(\theta)]}{\sqrt{\mathbb{E}_\theta[G_n(\theta)^2]}\sqrt{\mathbb{E}_\theta[U_n(\theta)^2]}} \tag{4.83}$$

is maximal for all $\theta \in \Theta$. Here we have used that

$$\mathbb{E}_\theta[U_n(\theta)] = \int \frac{\partial_\theta f(\,\cdot\,;\theta)}{f(\,\cdot\,;\theta)} f(\,\cdot\,;\theta) d\ell = \partial_\theta \int f(\,\cdot\,;\theta) d\ell = 0. \tag{4.84}$$

By a similar argument we obtain

$$\begin{aligned}
\mathbb{E}_\theta[G_n(\theta)U_n(\theta)] &= \int G_n(\theta)\partial_\theta f(\,\cdot\,;\theta) d\ell \\
&= \int \partial_\theta (G_n(\theta) f(\,\cdot\,;\theta)) d\ell \;-\; \int (\partial_\theta G_n(\theta)) f(\,\cdot\,;\theta) d\ell \\
&= \partial_\theta \underbrace{\mathbb{E}_\theta[G_n(\theta)]}_{=0} - \mathbb{E}_\theta[\partial_\theta G_n(\theta)] \\
&= -\mathbb{E}_\theta[\partial_\theta G_n(\theta)].
\end{aligned} \tag{4.85}$$

This implies that

$$\mathrm{Corr}_\theta[G_n, U_n]^2 = \frac{\left(\mathbb{E}_\theta[\partial_\theta G_n(\theta)]\right)^2}{\mathbb{E}_\theta[G_n(\theta)^2]\mathbb{E}_\theta[U_n(\theta)^2]}. \tag{4.86}$$

Thus we can maximize the correlation by minimizing $\mathbb{E}_\theta[G_n(\theta)^2]/(\mathbb{E}_\theta[\partial_\theta G_n(\theta)])^2$. If the correlation is negative, simply consider the estimating function $-G_n(\theta)$.

Definition 4.2.2.2. $G_n^* \in \mathcal{G}_n$ is called *F-optimal* in $\mathcal{G}_n$ if

$$\frac{\mathbb{E}_\theta[G_n(\theta)^2]}{(\mathbb{E}_\theta[\partial_\theta G_n(\theta)])^2} \geq \frac{\mathbb{E}_\theta[G_n^*(\theta)^2]}{(\mathbb{E}_\theta[\partial_\theta G_n^*(\theta)])^2}, \tag{4.87}$$

for all $\theta \in \Theta$ and for all $G_n \in \mathcal{G}_n$.

This type of optimality is called *F-optimality* since the sample size n was fixed in the preceding considerations. The F-optimal estimating function G_n^* is sometimes called the *quasi-score-function*, and the estimator obtained from it is called the *quasi-likelihood estimator*. The next theorem yields a criterion for F-optimality.

Theorem 4.2.2.3 (Godambe and Heyde, 1987)**.** *Suppose that $\mathcal{G}_n$ is closed under addition. If $G_n^* \in \mathcal{G}_n$ satisfies the inequality*

$$\mathbb{E}_\theta[(G_n(\theta) - U_n(\theta))^2] \geq \mathbb{E}_\theta[(G_n^*(\theta) - U_n(\theta))^2], \tag{4.88}$$

for all $\theta \in \Theta$ and for all $G_n \in \mathcal{G}_n$, then G_n^* is F-optimal in G_n.

Proof. See Godambe and Heyde [32]. □

Theorem 4.2.2.4 (Heyde, 1988)**.** *Suppose that $\mathcal{G}_n$ is closed under addition. Then $G_n^* \in \mathcal{G}_n$ is F-optimal in $\mathcal{G}_n$ if and only if*

$$\frac{\mathbb{E}_\theta[G_n(\theta)G_n^*(\theta)]}{\mathbb{E}_\theta[\partial_\theta G_n(\theta)]} = \frac{\mathbb{E}_\theta(G_n^*(\theta)^2)}{\mathbb{E}_\theta[\partial_\theta G_n^*(\theta)]}, \tag{4.89}$$

for all $\theta \in \Theta$ and for all $G_n \in \mathcal{G}_n$.

Proof. See Heyde [34]. □

Now we consider a different optimality concept based on the properties of the martingale estimating function and the resulting estimator as the sample size n tends to infinity.

For the Markov process X define the σ-algebra $\mathcal{F}_n = \sigma(X_s, s \leq n),\ n \in \mathbb{N}$. Recall that within this section, we consider the case where $\Delta = 1$ and that the stochastic process $G(\theta) = (G_n(\theta))_{n\in\mathbb{N}}$ is a martingale estimating function, since for every $\theta \in \Theta$, G is a $\mathbb{P}_\theta$ -martingale with respect to the filtration $\mathcal{F}_n$. We assumed that $G_n(\theta)$ is square integrable under $\mathbb{P}_\theta$. Thus the predictable quadratic variation process of $G(\theta)$ exists. We will denote this process with $\langle G(\theta)\rangle$.

As before, let θ_0 be the true parameter value and let $\hat\theta_n$ be an estimator obtained from G. A Taylor expansion yields

$$0 = G_n(\hat\theta_n) = G_n(\theta_0) + \partial_\theta G_n(\tilde\theta_n)(\hat\theta_n - \theta_0), \tag{4.90}$$

where $\tilde\theta_n$ is between $\hat\theta_n$ and θ_0. On certain additional assumptions, which we assume to be satisfied, the central limit theorem for martingales states that

$$\frac{G_n(\theta_0)}{\langle G(\theta_0)\rangle_n^{1/2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{4.91}$$

as $n \to \infty$. See e.g. Hall and Heyde [33]. This result and the following ones are of course under $\mathbb{P}_{\theta_0}$, since we assumed that θ_0 is the true parameter value. Together with (4.90) it follows that

$$\frac{\partial_\theta G_n(\tilde\theta)}{\langle G(\theta_0)\rangle_n^{1/2}}(\hat\theta_n - \theta_0) = -\frac{G_n(\theta_0)}{\langle G(\theta_0)\rangle_n^{1/2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{4.92}$$

as $n \to \infty$. Now suppose that

$$\hat\theta_n \longrightarrow \theta_0, \tag{4.93}$$

in probability as $n \to \infty$ and that

$$\frac{\partial_\theta G_n(\theta_0)}{\partial_\theta G_n(\tilde{\theta})} \longrightarrow 1, \tag{4.94}$$

in probability as $n \to \infty$. Conditions ensuring the existence of such an estimator were given in Section 4.2.1. On the assumptions made so far we have

$$\frac{\partial_\theta G_n(\theta_0)}{\langle G(\theta_0)\rangle_n^{1/2}}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N(0,1), \tag{4.95}$$

as $n \to \infty$. Usually $\partial_\theta G_n(\theta_0)$ is rather complicated. Often a simpler expression can be obtained in the following way. Let $\bar{G}_n(\theta_0)$ be the compensator of $\partial_\theta G_n(\theta_0)$ under $\mathbb{P}_{\theta_0}$. This means $\partial_\theta G_n(\theta_0) - \bar{G}_n(\theta_0)$ is a martingale with respect to $\mathcal{F}_n$ under $\mathbb{P}_{\theta_0}$, and $\bar{G}_n(\theta_0)$ is $\mathcal{F}_{n-1}$ measurable for all $n \in \mathbb{N}$. Suppose

$$\frac{\partial_\theta G_n(\theta_0)}{\bar{G}_n(\theta_0)} \longrightarrow 1 \tag{4.96}$$

in probability as $n \to \infty$. Then we eventually get

$$\frac{\bar{G}_n(\theta_0)}{\langle G(\theta_0)\rangle_n^{1/2}}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} N(0,1), \tag{4.97}$$

as $n \to \infty$. We see that the quantity

$$\frac{\langle G(\theta_0)\rangle_n}{\bar{G}_n(\theta_0)^2} \tag{4.98}$$

is a measure of the asymptotic random variation of $\hat{\theta}_n$ around the true parameter value θ_0. Thus we should try to minimize this quantity.

Let $\mathcal{G}$ be a class of unbiased martingale estimating functions satisfying the various regularity conditions imposed above. We introduce the following notion of *A-optimality*.

Definition 4.2.2.5. $G^* \in \mathcal{G}$ is called *A-optimal* in $\mathcal{G}$ if

$$\frac{\langle G(\theta)\rangle_n}{\bar{G}_n(\theta)^2} \geq \frac{\langle G^*(\theta)\rangle_n}{\bar{G}^*(\theta)^2}, \tag{4.99}$$

for all $n \in \mathbb{N}$, $\theta \in \Theta$ and $G \in \mathcal{G}$.

As we have mentioned above, $\langle G(\theta)\rangle_n$ denotes the predictable quadratic variation process of the martingale $G(\theta)$ under $\mathbb{P}_\theta$. *A-optimal* is short for asymptotically optimal. In the next theorem $\langle G(\theta), G^*(\theta)\rangle$ denotes the predictable quadratic covariation process of the martingales $G(\theta)$ and $G^*(\theta)$ under $\mathbb{P}_\theta$.

Theorem 4.2.2.6 (Heyde, 1988)**.** *Suppose $\mathcal{G}$ is closed under addition. Then $G^* \in \mathcal{G}$ is A-optimal in $\mathcal{G}$ if and only if*

$$\frac{\langle G(\theta), G^*(\theta)\rangle_n}{\bar{G}_n(\theta)} = \frac{\langle G^*(\theta)\rangle_n}{\bar{G}^*(\theta)} \tag{4.100}$$

for all $n \in \mathbb{N}$, $\theta \in \Theta$ and $G \in \mathcal{G}$.

Proof. See Heyde [34]. □

In order to study the relation between F-optimality and A-optimality, note that

$$\begin{aligned} \mathbb{E}_\theta[\langle G(\theta), G^*(\theta)\rangle_n] &= \mathbb{E}_\theta[G(\theta)G^*(\theta)], \\ \mathbb{E}_\theta[\bar{G}_n(\theta)] &= \mathbb{E}_\theta[\partial_\theta G_n(\theta)]. \end{aligned}$$

The following result holds.

Theorem 4.2.2.7 (Heyde, 1988)**.** *Suppose that $\mathcal{G}$ is closed under addition, that $G^* \in \mathcal{G}$ is A-optimal in $\mathcal{G}$ and that $\eta_n(\theta) = \langle G^*(\theta)\rangle_n / \bar{G}_n^*(\theta)$ is non-random for some $n \in \mathbb{N}$. Then $G_n^*(\theta)$ is F-optimal in $\mathcal{G}_n = \{G_n : G \in \mathcal{G}\}$.*

Proof. See Heyde [34]. □

4.2.3 A- and F-optimal generalized martingale estimating functions

Having introduced the concept of A- and F-optimality we wish to answer the following question: When is an estimating function of the type

$$G_n^{(H,L,X)}(\theta) = \sum_{i=1}^{n} g(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i}) \tag{4.101}$$

optimal? For fixed $\Delta > 0$, θ and x, let

$$(h, l, y) \longmapsto f_{(H,L,X)}(\Delta, x, h, l, y; \theta) = f(\Delta, x, h, l, y; \theta) \tag{4.102}$$

denote the density of $(H_\Delta, L_\Delta, X_\Delta)$, conditional on $X_0 = x$, and let $E(x) = \{(h, l, y) \in \mathbb{R}^3 \,|\, l \leq x, y \leq h\}$ denote the state space of the triplet $(H_\Delta, L_\Delta, X_\Delta)$. We consider the Hilbert space

$$L^2\Big(E(x), f(t, x, h, l, y; \theta) dh\, dl\, dy\Big) = L^2(E, f, x). \tag{4.103}$$

The scalar product on $L^2(E, f, x)$ is defined as

$$\langle u, v\rangle_{L^2(E,f,x)} = \int_{E(x)} u(h, l, y) v(h, l, y) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy. \tag{4.104}$$

The $L^2(E,f,x)$-norm is given by

$$\|u\|^2_{L^2(E,f,x)} = \int_{E(x)} |u(h,l,y)|^2 f(\Delta,x,h,l,y;\theta)dh\,dl\,dy. \tag{4.105}$$

Moreover, we define a subclass of functions $K(\Delta,x,h,l,y;\theta) \subset L^2(E,f,x)$ that consists of functions of the following type:

$$\sum_{j=1}^{N} \beta_j k_j(\Delta,x,h,l,y;\theta), \tag{4.106}$$

where $\beta_j \in \mathbb{R}$ and $N \in \mathbb{N}$ is fixed. We formulate two conditions.

Condition 4.2.3.1.

$$\int_{E(x)} k_j(\Delta,x,h,l,y;\theta)^2 f(\Delta,x,h,l,y;\theta)dh\,dl\,dy < \infty, \tag{4.107}$$

for all $\Delta > 0,\ x \in \mathbb{R},\ \theta \in \Theta$ *and* $j = 1,\dots,N$.

Condition 4.2.3.2.

1. $f(\Delta,x,h,l,y;\theta)$ *is differentiable w.r.t.* θ*, for all* Δ, x *and for all* $h,l,y \in E(x)$.
2. $\partial_\theta f(\Delta,x,h,l,y;\theta) \in L^2(E,f,x)$ *for all* Δ,x,θ *and for all* $h,l,y \in E(x)$.
3. $(h,l,y,\theta) \longmapsto \partial_\theta[k_j(\Delta,x,h,l,y;\theta)f(\Delta,x,h,l,y;\theta)]$ *is locally dominated integrable with respect to the Lebesgue measure for every* Δ, x *and* $h,l,y \in E(x)$ *and for all* $j = 1,\dots,N$.
4. $k_j(\Delta,x,h,l,y;\theta)$, $j = 1,\dots,N$, *are linearly independent in* $L^2(E,f,x)$.

With these conditions at hand, we are able to state the main result of this paragraph.

Theorem 4.2.3.3. *Suppose that Condition 4.2.3.1 and Condition 4.2.3.2 are satisfied and let* $g^*(\Delta,x,h,l,y;\theta)$ *denote the orthogonal projection of* $\partial_\theta \log f(\Delta,x,h,l,y;\theta)$ *onto* $K(\Delta,x,h,l,y;\theta)$ *with respect to the scalar product* $\langle\cdot,\cdot\rangle_{L^2(E,f,x)}$*. Furthermore, suppose that* $g^*(\Delta,x,h,l,y;\theta)$ *is continuously differentiable with respect to* θ *for all* Δ, x *and for all* $h,l,y \in E(x)$*. Define the estimating function* G^* *by*

$$G_n^*(\theta) = \sum_{i=1}^{N} g^*(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i}). \tag{4.108}$$

Then G^* *is A-optimal in* $\mathcal{G}$ *and* G_n^* *is F-optimal in* $\mathcal{G}_n$ *for all* $n \in \mathbb{N}$.

Before we prove Theorem 4.2.3.3, we want to calculate the function g^*. It can be written as

$$g^*(\Delta,x,h,l,y;\theta) = \sum_{j=1}^{N} a_j^*(\Delta,x;\theta)k_j(\Delta,x,h,l,y;\theta). \tag{4.109}$$

Hence, determining g^* is tantamount to determining the optimal weights $(a_1^*, \ldots, a_N^*)$. Since g^* is the orthogonal projection of $\partial_\theta \log f(\Delta, x, h, l, y; \theta)$ onto $K(\Delta, x, h, l, y; \theta)$ with respect to $\langle \cdot, \cdot \rangle_{L^2(E,f,x)}$ the real numbers $a_j^*(\Delta, x; \theta)$ can be found by solving the equations

$$\big\langle \partial_\theta \log f - g^*, k_l \big\rangle_{L^2(E,f,x)} = 0, \quad l = 1, \ldots, N. \tag{4.110}$$

These equations can be reformulated in the following form

$$\big\langle \partial_\theta \log f, k_l \big\rangle_{L^2(E,f,x)} = \sum_{j=1}^{N} a_j^*(\Delta, x; \theta) \big\langle k_j, k_l \big\rangle_{L^2(E,f,x)}, \quad l = 1, \ldots, N. \tag{4.111}$$

For $l = 1, \ldots, N$, define the real numbers

$$\begin{aligned} b_l(\Delta, x; \theta) &= \big\langle \partial_\theta \log f, k_l \big\rangle_{L^2(E,f,x)} \\ &= \int_{E(x)} k_l(\Delta, x, h, l, y; \theta) \frac{\partial_\theta f(\Delta, x, h, l, y; \theta)}{f(\Delta, x, h, l, y; \theta)} f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &= \int_{E(x)} \partial_\theta \big[k_l(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta)\big] dh\, dl\, dy \\ &\quad - \int_{E(x)} \big[\partial_\theta k_l(\Delta, x, h, l, y; \theta)\big] f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &= \partial_\theta \int_{E(x)} k_l(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &\quad - \int_{E(x)} \big[\partial_\theta k_l(\Delta, x, h, l, y; \theta)\big] f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &= - \int_{E(x)} \big[\partial_\theta k_l(\Delta, x, h, l, y; \theta)\big] f(\Delta, x, h, l, y; \theta)\ dh\, dl\, dy, \end{aligned} \tag{4.112}$$

and for $j, l = 1, \ldots, N$, define

$$\begin{aligned} c_{jl}(\Delta, x; \theta) &= \langle k_j, k_l \rangle_{L^2(E,f,x)} \\ &= \int_{E(x)} k_j(\Delta, x, h, l, y; \theta) k_l(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta)\ dh\, dl\, dy. \end{aligned} \tag{4.113}$$

Introducing the $N \times N$-matrix

$$C(\Delta, x; \theta) = \Big(c_{jl}(\Delta, x; \theta)\Big)_{j,l=1,\ldots,N} \tag{4.114}$$

and the following two row vectors

$$B(\Delta, x; \theta) = (b_1(\Delta, x; \theta), \ldots, b_N(\Delta, x; \theta)) \tag{4.115}$$

and

$$A^*(\Delta, x; \theta) = (a_1^*(\Delta, x; \theta), \ldots, a_N^*(\Delta, x; \theta)), \tag{4.116}$$

we are able to rewrite the equations (4.111) in the following way

$$B(\Delta, x; \theta) = A^*(\Delta, x; \theta) C(\Delta, x; \theta). \tag{4.117}$$

On the Condition 4.2.3.2 the matrix $C(\Delta, x; \theta)$ is invertible since the functions k_j, $j = 1, \ldots, N$, were assumed to be linearly independent in $L^2(E, f, x)$. Thus (4.117) becomes

$$A^*(\Delta, x; \theta) = B(\Delta, x; \theta)\, C(\Delta, x; \theta)^{-1}. \tag{4.118}$$

Altogether, we obtain

$$\begin{aligned} g^*(\Delta, x, h, l, y; \theta) &= A^*(\Delta, x; \theta) k(\Delta, x, h, l, y; \theta) \\ &= B(\Delta, x; \theta) C(\Delta, x; \theta)^{-1} k(\Delta, x, h, l, y; \theta), \end{aligned} \tag{4.119}$$

where $k = (k_1, \ldots, k_N)^T$. We are now able to prove Theorem 4.2.3.3.

Proof (of Theorem 4.2.3.3). Define the two row vectors

$$A(\Delta, x; \theta) = (a_j(\Delta, x; \theta))_{j=1,\ldots,N} \tag{4.120}$$

and

$$\widetilde{A}(\Delta, x; \theta) = (\tilde{a}_j(\Delta, x; \theta))_{j=1,\ldots,N}\,, \tag{4.121}$$

and the column vector

$$k(\Delta, x, h, l, y; \theta) = \begin{pmatrix} k_1(\Delta, x, h, l, y; \theta) \\ \vdots \\ k_N(\Delta, x, h, l, y; \theta) \end{pmatrix}. \tag{4.122}$$

Then we are able to define the two estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)}; \theta) k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i}; \theta) \tag{4.123}$$

and

$$\widetilde{G}_n(\theta) = \sum_{i=1}^{n} \widetilde{A}(\Delta, X_{\Delta(i-1)}; \theta) k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i}; \theta). \tag{4.124}$$

The covariation process of $G_n(\theta)$ and $\tilde{G}_n(\theta)$ is given by

$$
\begin{aligned}
&\langle G(\theta), \widetilde{G}(\theta)\rangle_n \\
&\quad = \sum_{i=1}^{n} \mathbb{E}_\theta\big[A(\Delta, X_{\Delta(i-1)};\theta) k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta) \\
&\qquad\qquad \times\ k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta)^T \widetilde{A}(\Delta, X_{\Delta(i-1)};\theta)^T \big| \mathcal{F}_{\Delta(i-1)}\big] \\
&\quad = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta) \\
&\qquad\qquad \times \mathbb{E}_\theta\big[k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta) k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta)^T \big| X_{\Delta(i-1)}\big] \\
&\qquad\qquad \times\ \widetilde{A}(\Delta, X_{\Delta(i-1)};\theta)^T \\
&\quad = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta) \widetilde{A}(\Delta, X_{\Delta(i-1)};\theta)^T, \qquad (4.125)
\end{aligned}
$$

where the entries $c_{jl}(\Delta, x;\theta)$ of the matrix $C(\Delta, x;\theta)$ are given by

$$
c_{jl}(\Delta, x;\theta) = \langle k_j, k_l\rangle_{L^2(E,f,x)}, \quad j,l = 1,\ldots,N. \qquad (4.126)
$$

Particularly, we have

$$
\begin{aligned}
&\langle G^*(\theta)\rangle_n \\
&\quad = \sum_{i=1}^{n} A^*(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta) A^*(\Delta, X_{\Delta(i-1)};\theta)^T \\
&\quad = \sum_{i=1}^{n} B(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta)^{-1} C(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta)^{-1} \\
&\qquad\qquad \times B(\Delta, X_{\Delta(i-1)};\theta)^T \\
&\quad = \sum_{i=1}^{n} B(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta) B(\Delta, X_{\Delta(i-1)};\theta)^T \qquad (4.127)
\end{aligned}
$$

and the covariation process of $G(\theta)$ and $G^*(\theta)$ becomes

$$
\begin{aligned}
&\langle G(\theta), G^*(\theta)\rangle_n \\
&\quad = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta) A^*(\Delta, X_{\Delta(i-1)};\theta)^T \\
&\quad = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta) C(\Delta, X_{\Delta(i-1)};\theta)^{-1} B(\Delta, X_{\Delta(i-1)};\theta)^T \\
&\quad = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta) B(\Delta, X_{\Delta(i-1)};\theta)^T . \qquad (4.128)
\end{aligned}
$$

Recall that $\bar{G}_n(\theta)$ denotes the compensator of $\partial_\theta G_n(\theta)$. In order to determine $\bar{G}_n(\theta)$, we differentiate $G_n(\theta)$ with respect to θ:

$$\partial_\theta G_n(\theta) = \underbrace{\sum_{i=1}^{n} \partial_\theta A(\Delta, X_{\Delta(i-1)};\theta) k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta)}_{\mathbb{P}_\theta-\text{martingale}}$$
$$+ \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta)\partial_\theta k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta). \tag{4.129}$$

Hence, $\bar{G}_n(\theta)$ is the compensator of the second addend on the right hand side of the previous equation. This means

$$\bar{G}_n(\theta) = \sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta)\mathbb{E}_\theta\left[\partial_\theta k(\Delta, X_{\Delta(i-1)}, H_{\Delta i}, L_{\Delta i}, X_{\Delta i};\theta)\middle| X_{\Delta(i-1)}\right], \tag{4.130}$$

or equivalently

$$\bar{G}_n(\theta) = -\sum_{i=1}^{n} A(\Delta, X_{\Delta(i-1)};\theta)B(\Delta, X_{\Delta(i-1)};\theta)^T = -\langle G(\theta), G^*(\theta)\rangle_n. \tag{4.131}$$

Especially, we conclude that

$$\bar{G}^*_n(\theta) = -\langle G^*(\theta)\rangle_n. \tag{4.132}$$

The assertion now follows directly from Theorems 4.2.2.6 and 4.2.2.7, since the quantity

$$\bar{G}_n(\theta)^{-1}\langle G(\theta), G^*(\theta)\rangle_n = -1 = \bar{G}^*_n(\theta)^{-1}\langle G^*(\theta)\rangle_n \tag{4.133}$$

is non-random. □

Asymptotic Variance in the A-optimal Case

In Theorem 4.2.1.7 we stated that - on suitable assumptions - the estimator $\hat{\theta}_n$ is asymptotically normally distributed. Recall that, under the true measure $\mathbb{P}_{\theta_0}$, we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \longrightarrow N\left(0, \frac{v(\theta_0)}{\xi^2(\theta_0)}\right), \tag{4.134}$$

weakly as $n \to \infty$. Here, $v(\theta_0) = Q^\Delta_{\theta_0}(g(\Delta, \cdot;\theta_0)^2)$ and $\xi(\theta_0) = Q^\Delta_{\theta_0}(\partial_\theta g(\Delta, \cdot;\theta_0))$.

The function ξ clearly satisfies

$$\begin{aligned}\xi(\theta_0) &= Q^\Delta_{\theta_0}(\partial_\theta g(\Delta, x, h, l, y;\theta_0)) \\ &= Q^\Delta_{\theta_0}\left(\sum_{j=1}^{N} \partial_\theta a_j(\Delta, x;\theta)k_j(\Delta, x, h, l, y;\theta) + a_j(\Delta, x;\theta)\partial_\theta k_j(\Delta, x, h, l, y;\theta)\right)\end{aligned}$$

$$
\begin{aligned}
&= Q^{\Delta}_{\theta_0}\left(\sum_{j=1}^{N} a_j(\Delta,x;\theta)\partial_\theta k_j(\Delta,x,h,l,y;\theta)\right) \\
&= \sum_{j=1}^{N} Q^{\Delta}_{\theta_0}\Big(a_j(\Delta,x;\theta)\partial_\theta k_j(\Delta,x,h,l,y;\theta)\Big). \qquad (4.135)
\end{aligned}
$$

If we insert the optimal weights a^* given by (4.118), the latter equation becomes

$$
\begin{aligned}
\xi(\theta_0) &= \sum_{j=1}^{N} Q^{\Delta}_{\theta_0}(a^*_j(\Delta,x;\theta)\partial_\theta k_j(\Delta,x,h,l,y;\theta)) \\
&= -\sum_{j=1}^{N} Q^{\Delta}_{\theta_0}\left(a^*_j(\Delta,x;\theta)\sum_{k=1}^{N} k_j(\Delta,x,h,l,y;\theta)k_j(\Delta,x,h,l,y;\theta)a^*_k(\Delta,x;\theta)\right) \\
&= -Q^{\Delta}_{\theta_0}(g^*(\Delta,x,h,l,y;\theta)^2) = -v(\theta_0). \qquad (4.136)
\end{aligned}
$$

Note that, for the second equality, we made use of (4.117) and thus the asymptotic variance of an A-optimal estimators is $1/v(\theta_0)$. All in all, on the assumptions made in Theorem 4.2.1.7, there is a consistent and A-optimal estimator $\hat{\theta}^*_n$ that satisfies

$$
\sqrt{n}(\hat{\theta}^*_n - \theta_0) \longrightarrow N\left(0, \frac{1}{v(\theta_0)}\right), \qquad (4.137)
$$

weakly as $n \to \infty$.

4.3 Concrete examples

4.3.1 The optimal linear estimator

Let us go back to the example of a linear estimating function which is defined by (4.23) and (4.24) for $N = 3$ with

$$
\begin{aligned}
k_1(\Delta,x,h,l,y;\theta) &= h - F^H(\Delta,x;\theta), \\
k_2(\Delta,x,h,l,y;\theta) &= l - F^L(\Delta,x;\theta), \\
k_3(\Delta,x,h,l,y;\theta) &= y - F^X(\Delta,x;\theta), \qquad (4.138)
\end{aligned}
$$

where

$$
F^H(\Delta,x;\theta) = \mathbb{E}_{x,\theta}[\,H_\Delta\,], \quad F^L(\Delta,x;\theta) = \mathbb{E}_{x,\theta}[\,L_\Delta\,] \qquad (4.139)
$$

and

$$
F^X(\Delta,x;\theta) = \mathbb{E}_{x,\theta}[\,X_\Delta\,]. \qquad (4.140)
$$

In the notation of the previous paragraph, the diagonal entries of the matrix

$$C(\Delta, x; \theta) = \big(c_{jl}(\Delta, x; \theta)\big)_{j,l=1,2,3} = \big(\langle k_j, k_l \rangle\big)_{j,l=1,2,3} \tag{4.141}$$

become

$$\begin{aligned}
c_{11}(\Delta, x; \theta) &= \int_{E(x)} (h - F^H(\Delta, x; \theta))^2 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy = \phi_{H,H}(\Delta, x; \theta),\\
c_{22}(\Delta, x; \theta) &= \int_{E(x)} (l - F^L(\Delta, x; \theta))^2 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy = \phi_{L,L}(\Delta, x; \theta),\\
c_{33}(\Delta, x; \theta) &= \int_{E(x)} (y - F^X(\Delta, x; \theta))^2 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy = \phi_{X,X}(\Delta, x; \theta),
\end{aligned} \tag{4.142}$$

where

$$\phi_{U,V}(\Delta, x; \theta) = \mathrm{Cov}_{x,\theta}[U, V], \quad \text{for } U, V \in \{X_\Delta, H_\Delta, L_\Delta\}. \tag{4.143}$$

Moreover, we have the following off-diagonal entries

$$\begin{aligned}
c_{12}(\Delta, x; \theta) &= \int_{E(x)} (h - F^H(\Delta, x; \theta))(l - F^L(\Delta, x; \theta)) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy\\
&= \phi_{H,L}(\Delta, x; \theta),\\
c_{13}(\Delta, x; \theta) &= \int_{E(x)} (h - F^H(\Delta, x; \theta))(y - F^X(\Delta, x; \theta)) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy\\
&= \phi_{H,X}(\Delta, x; \theta),\\
c_{23}(\Delta, x; \theta) &= \int_{E(x)} (L - F^L(\Delta, x; \theta))(y - F^X(\Delta, x; \theta)) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy\\
&= \phi_{L,X}(\Delta, x; \theta).
\end{aligned} \tag{4.144}$$

All in all, the matrix $C(\Delta, x; \theta)$ becomes

$$C(\Delta, x; \theta) = \begin{pmatrix} \phi_{H,H}(\Delta, x; \theta) & \phi_{H,L}(\Delta, x; \theta) & \phi_{H,X}(\Delta, x; \theta) \\ \phi_{L,H}(\Delta, x; \theta) & \phi_{L,L}(\Delta, x; \theta) & \phi_{L,X}(\Delta, x; \theta) \\ \phi_{X,H}(\Delta, x; \theta) & \phi_{X,L}(\Delta, x; \theta) & \phi_{X,X}(\Delta, x; \theta) \end{pmatrix}. \tag{4.145}$$

Its determinant is

$$\begin{aligned}
&\det\big(C(\Delta, x; \theta)\big)\\
&= -\phi_{H,X}(\Delta, x; \theta)^2 \phi_{LL}(\Delta, x; \theta) + 2\phi_{H,L}(\Delta, x; \theta)\phi_{H,X}(\Delta, x; \theta)\phi_{L,X}(\Delta, x; \theta)\\
&\quad - \phi_{H,L}(\Delta, x; \theta)^2 \phi_{X,X}(\Delta, x; \theta) - \phi_{L,X}(\Delta, x; \theta)^2 \phi_{H,H}(\Delta, x; \theta)^2\\
&\quad + \phi_{L,L}(\Delta, x; \theta)\phi_{X,X}(\Delta, x; \theta)\phi_{H,H}(\Delta, x; \theta)
\end{aligned} \tag{4.146}$$

and consequently the expression $\det\left(C(\Delta, x; \theta)\right)\ C(\Delta, x; \theta)^{-1}$ equals

$$\begin{pmatrix} -\phi_{L,X}^2 + \phi_{L,L}\phi_{X,X} & \phi_{H,X}\phi_{L,X} - \phi_{H,L}\phi_{X,X} & -\phi_{H,X}\phi_{L,L} + \phi_{H,L}\phi_{L,X} \\ \phi_{H,X}\phi_{L,X} - \phi_{H,L}\phi_{X,X} & -\phi_{H,X}^2 + \phi_{H,H}\phi_{X,X} & \phi_{H,L}\phi_{H,X} - \phi_{L,X}\phi_{H,H} \\ -\phi_{H,X}\phi_{L,L} + \phi_{H,L}\phi_{L,X} & \phi_{H,L}\phi_{H,X} - \phi_{L,X}\phi_{H,H} & -\phi_{H,L}^2 + \phi_{L,L}\phi_{H,H} \end{pmatrix} (\Delta, x; \theta). \tag{4.147}$$

The vector $B(\Delta, x; \theta)$ consists of the entries

$$b_j(\Delta, x; \theta) = -\int_{E(x)} \partial_\theta k_j(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy,\ \ j = 1, 2, 3. \tag{4.148}$$

Therefore,

$$\begin{aligned} b_1(\Delta, x; \theta) &= -\int_{E(x)} \partial_\theta k_1(\Delta, x, h, l, y; \theta) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &= -\int_{E(x)} \partial_\theta (h - F^H(\Delta, x; \theta)) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &= \partial_\theta F^H(\Delta, x; \theta), \end{aligned} \tag{4.149}$$

and analogously

$$b_2(\Delta, x; \theta) = \partial_\theta F^L(\Delta, x; \theta) \quad \text{and} \quad b_3(\Delta, x; \theta) = \partial_\theta F^X(\Delta, x; \theta). \tag{4.150}$$

Overall, we get the optimal weights

$$A^*(\Delta, x; \theta) = (a_1^*(\Delta, x; \theta), a_2^*(\Delta, x; \theta), a_3^*(\Delta, x; \theta))^T \tag{4.151}$$

with

$$\begin{aligned} &a_1^*(\Delta, x; \theta) \\ &= \det\left(C(\Delta, x; \theta)\right)^{-1} \Big\{ (-\phi_{L,X}^2 + \phi_{L,L}\phi_{X,X}) \partial_\theta F^H + (\phi_{H,X}\phi_{L,X} - \phi_{H,L}\phi_{X,X}) \partial_\theta F^L \\ &\qquad + (-\phi_{H,X}\phi_{L,L} + \phi_{H,L}\phi_{L,X}) \partial_\theta F^X \Big\} (\Delta, x; \theta), \end{aligned} \tag{4.152}$$

$$\begin{aligned} &a_2^*(\Delta, x; \theta) \\ &= \det\left(C(\Delta, x; \theta)\right)^{-1} \Big\{ (\phi_{H,X}\phi_{L,X} - \phi_{H,L}\phi_{X,X}) \partial_\theta F^H + (-\phi_{H,X}^2 + \phi_{H,H}\phi_{X,X}) \partial_\theta F^L \\ &\qquad + (\phi_{H,L}\phi_{H,X} - \phi_{L,X}\phi_{H,H}) \partial_\theta F^X \Big\} (\Delta, x; \theta) \end{aligned} \tag{4.153}$$

and

$$\begin{aligned} &a_3^*(\Delta, x; \theta) \\ &= \det\left(C(\Delta, x; \theta)\right)^{-1} \Big\{ (-\phi_{H,X}\phi_{L,L} + \phi_{H,L}\phi_{L,X}) \partial_\theta F^H + (\phi_{H,L}\phi_{H,X} - \phi_{L,X}\phi_{H,H}) \partial_\theta F^L \end{aligned}$$

$$+ (-\phi_{H,L}^2 + \phi_{L,L}\phi_{H,H})\partial_\theta F^X \Big\}(\Delta, x; \theta). \tag{4.154}$$

4.3.2 Optimal quadratic estimators

The most general quadratic estimating function that we are able to consider in our model is given for $N = 9$ by the functions $k_i,\ i = 1, \ldots, 9$, defined by

$$\begin{aligned} k_1(\Delta, x, h, l, y; \theta) &= h - F^H(\Delta, x; \theta), \\ k_2(\Delta, x, h, l, y; \theta) &= l - F^L(\Delta, x; \theta), \\ k_3(\Delta, x, h, l, y; \theta) &= y - F^X(\Delta, x; \theta), \end{aligned} \tag{4.155}$$

and

$$\begin{aligned} k_4(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)][l - F^L(\Delta, x; \theta)] - \phi_{H,L}(\Delta, x; \theta), \\ k_5(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)][y - F^X(\Delta, x; \theta)] - \phi_{H,X}(\Delta, x; \theta), \\ k_6(\Delta, x, h, l, y; \theta) &= [l - F^L(\Delta, x; \theta)][y - F^X(\Delta, x; \theta)] - \phi_{L,X}(\Delta, x; \theta), \\ k_7(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)]^2 - \phi_{H,H}(\Delta, x; \theta), \\ k_8(\Delta, x, h, l, y; \theta) &= [l - F^L(\Delta, x; \theta)]^2 - \phi_{L,L}(\Delta, x; \theta), \\ k_9(\Delta, x, h, l, y; \theta) &= [y - F^X(\Delta, x; \theta)]^2 - \phi_{X,X}(\Delta, x; \theta). \end{aligned} \tag{4.156}$$

Here, we set

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[\,U\,], \quad \text{for } U \in \{X_\Delta, H_\Delta, L_\Delta\}, \tag{4.157}$$

and

$$\phi_{U,V}(\Delta, x; \theta) = \mathrm{Cov}_{x,\theta}[\,U\,,\,V\,], \quad \text{for } U, V \in \{X_\Delta, H_\Delta, L_\Delta\}. \tag{4.158}$$

We are not going to display the entries of the matrix

$$C(\Delta, x; \theta) = \big(c_{jl}(\Delta, x; \theta)\big)_{j,l=1,\ldots,9} = \big(\langle k_j, k_l \rangle\big)_{j,l=1,\ldots,9}, \tag{4.159}$$

since this is not very demonstrative. Instead, we analyze a special case of the quadratic estimating function. Let $N = 3$ and define the functions

$$\begin{aligned} \widetilde{k}_1(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)]^2 - \phi_{H,H}(\Delta, x; \theta), \\ \widetilde{k}_2(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)][y - F^X(\Delta, x; \theta)] - \phi_{H,X}(\Delta, x; \theta), \\ \widetilde{k}_3(\Delta, x, h, l, y; \theta) &= [y - F^X(\Delta, x; \theta)]^2 - \phi_{X,X}(\Delta, x; \theta). \end{aligned} \tag{4.160}$$

Note that we will reconsider this particular estimating function and the resulting estimator in Chapter 6 when we examine small-Δ-optimal martingale estimating functions.

In the present model, the matrix

$$\widetilde{C}(\Delta, x; \theta) = \left(\widetilde{c}_{lj}(\Delta, x; \theta)\right)_{l,j=1,\ldots,3} = \left(\langle \widetilde{k}_l, \widetilde{k}_j \rangle\right)_{l,j=1,\ldots,3} \tag{4.161}$$

consists of the following entries. First, the diagonal entries are given by

$$\begin{aligned}\widetilde{c}_{11}(\Delta, x; \theta) = \int_{E(x)} &(h - F^H(\Delta, x; \theta))^4 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &- \phi_{H,H}(\Delta, x; \theta)^2,\end{aligned} \tag{4.162}$$

$$\begin{aligned}\widetilde{c}_{22}(\Delta, x; \theta) = \int_{E(x)} &(h - F^H(\Delta, x; \theta))^2 (y - F^X(\Delta, x; \theta))^2 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &- \phi_{H,X}(\Delta, x; \theta)^2,\end{aligned} \tag{4.163}$$

and

$$\widetilde{c}_{33}(\Delta, x; \theta) = \int_{E(x)} (y - F^X(\Delta, x; \theta))^4 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy - \phi_{X,X}(\Delta, x; \theta)^2. \tag{4.164}$$

The off-diagonal entries are given by

$$\begin{aligned}\widetilde{c}_{12}(\Delta, x; \theta) = \int_{E(x)} &(h - F^H(\Delta, x; \theta))^3 (y - F^X(\Delta, x; \theta)) f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &- \phi_{H,H}(\Delta, x; \theta)\phi_{H,X}(\Delta, x; \theta),\end{aligned} \tag{4.165}$$

$$\begin{aligned}\widetilde{c}_{13}(\Delta, x; \theta) = \int_{E(x)} &(h - F^H(\Delta, x; \theta))^2 (y - F^X(\Delta, x; \theta))^2 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &- \phi_{H,H}(\Delta, x; \theta)\phi_{X,X}(\Delta, x; \theta),\end{aligned} \tag{4.166}$$

and

$$\begin{aligned}\widetilde{c}_{23}(\Delta, x; \theta) = \int_{E(x)} &(h - F^H(\Delta, x; \theta))(y - F^X(\Delta, x; \theta))^3 f(\Delta, x, h, l, y; \theta) dh\, dl\, dy \\ &- \phi_{H,X}(\Delta, x; \theta)\phi_{X,X}(\Delta, x; \theta).\end{aligned} \tag{4.167}$$

For convenience, we generalize our previous notation. Henceforth, we write

$$\begin{aligned}&\widetilde{\phi}_{H^a,X^b}(\Delta, x; \theta) \\ &= \int_{E(x)} (h - F^H(\Delta, x; \theta))^a (y - F^X(\Delta, x; \theta))^b f(\Delta, x, h, l, y; \theta) dh\, dl\, dy,\end{aligned} \tag{4.168}$$

for $a, b \in \mathbb{N}_0$. Then, according to our earlier definitions, we have

$$\widetilde{\phi}_{H^{2a},1}(\Delta, x; \theta) = \phi_{H^a,H^a}(\Delta, x; \theta),$$

$$\begin{aligned}
\widetilde{\phi}_{H^a,X^b}(\Delta, x; \theta) &= \phi_{H^a,X^b}(\Delta, x; \theta), \\
\widetilde{\phi}_{1,X^{2a}}(\Delta, x; \theta) &= \phi_{X^a,X^a}(\Delta, x; \theta),
\end{aligned} \tag{4.169}$$

and consequently the entries of the symmetric matrix $\widetilde{C}(\Delta, x; \theta)$ are given by

$$\begin{aligned}
\widetilde{C}(\Delta, x; \theta)_{1,1} &= \widetilde{\phi}_{H^4,1}(\Delta, x; \theta) - \phi_{H,H}(\Delta, x; \theta)^2 \\
\widetilde{C}(\Delta, x; \theta)_{1,2} = \widetilde{C}(\Delta, x; \theta)_{2,1} &= \widetilde{\phi}_{H^3,X}(\Delta, x; \theta) - \phi_{H,H}(\Delta, x; \theta)\phi_{H,X}(\Delta, x; \theta) \\
\widetilde{C}(\Delta, x; \theta)_{1,3} = \widetilde{C}(\Delta, x; \theta)_{3,1} &= \widetilde{\phi}_{H^2,X^2}(\Delta, x; \theta) - \phi_{H,H}(\Delta, x; \theta)\phi_{X,X}(\Delta, x; \theta) \\
\widetilde{C}(\Delta, x; \theta)_{2,2} &= \widetilde{\phi}_{H^2,X^2}(\Delta, x; \theta) - \phi_{H,X}(\Delta, x; \theta)^2 \\
\widetilde{C}(\Delta, x; \theta)_{2,3} = \widetilde{C}(\Delta, x; \theta)_{3,2} &= \widetilde{\phi}_{H,X^3}(\Delta, x; \theta) - \phi_{H,X}(\Delta, x; \theta)\phi_{X,X}(\Delta, x; \theta) \\
\widetilde{C}(\Delta, x; \theta)_{3,3} &= \widetilde{\phi}_{1,X^4}(\Delta, x; \theta) - \phi_{X,X}(\Delta, x; \theta)^2
\end{aligned} \tag{4.170}$$

The calculation of the determinant of $\widetilde{C}(\Delta, x; \theta)$ and the optimal weights $a_j^*(\Delta, x; \theta),\ j = 1, 2, 3$, now is straightforward. Yet, the computations are tedious and therefore omitted.

5 Second Order Expansions

5.1 Introduction & Motivation

In this chapter we consider time-homogeneous diffusion processes X on $\mathbb{R}$. For an introduction to diffusions, see Chapter 2. From X the running maximum and the running minimum can be inferred. We denote these processes with H and L, respectively. In Chapter 4 we presented results for martingale estimating functions in a parameterized model. For a fixed sampling frequency $\Delta > 0$ these estimating functions are based on the knowledge of the joint density $f_{(H,L,X)}(\Delta, x, h, l, y)$ of $(H_\Delta, L_\Delta, X_\Delta)$, conditional on $X_0 = x$. In Chapter 3 we saw that such a density, if it exists, can be calculated by means of the transition density $p^{(l,h)}(\Delta, x, y)$ of the diffusion X killed at the boundary of the interval (l, h), $l < h$. Unfortunately, we usually do not know $p^{(l,h)}$ explicitly. As a result, extensive simulations of the diffusion X or numerical methods for partial differential equations are necessary to find an approximation to the joint density of $(H_\Delta, L_\Delta, X_\Delta)$. However, both numerical methods require an enormous computational effort. This is the main motivation for the search of simplified inference methods. This chapter establishes the foundations for the analysis of approximately optimal martingale estimating functions constructed from approximations of the moments of the triplet (H, L, X).

More precisely, the aim of the present chapter is to find an expansion of the expression $\mathbb{E}_x[g(H_t, L_t, X_t)]$, with respect to the square root of the time variable t, and for sufficiently smooth functions $g : \mathbb{R}^3 \to \mathbb{R}$ that do not grow too fast. Our approach relies entirely on elementary estimates like Doob's maximal inequality and Cauchy-Schwarz' inequality. The result is an expansion whose highest order term is proportional to $\sqrt{t}^2 = t$. Higher order expansions cannot be derived by this method. However, as we will see in Chapter 6, the presented second order expansion, with respect to $\sqrt{t}$, suffices to analyze the asymptotical behavior of martingale estimating functions constructed from a fixed-size sample as the sampling frequency $\Delta = t$ tends to 0. Finally, let us note that a completely different approach will put us into a position to derive higher order expansions of $\mathbb{E}_x[g(H_t, X_t)]$ as well. The shortcoming is that, in order to calculate higher order terms, more advanced, and hence, more difficult techniques are required. For more details, see Chapter 7 and Chapter 8. Throughout the present chapter, we will stick to the notations of Chapter 2.

5.2 Second Order Expansions

5.2.1 Auxiliary Results

First, we have to state auxiliary results that will turn out to be crucial for finding a second order expansion of $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$. Let B denote the standard Brownian motion of $\mathbb{R}$ and let X be a diffusion process that satisfies

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x, \ t \geq 0, \tag{5.1}$$

with sufficiently smooth coefficients μ and σ. Usually, we will postulate the following minimal condition.

Condition 5.2.1.1. *There is a constant $K > 0$ such that the coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ satisfy the following Lipschitz condition*

$$\left|\mu(x) - \mu(y)\right| + \left|\sigma(x) - \sigma(y)\right| \leq K|x - y|, \tag{5.2}$$

for all $x, y \in \mathbb{R}$. Note that this particularly implies a linear growth condition for both coefficients. Eventually, the coefficient σ is supposed to be uniformly bounded away from zero.

Let us state our first result.

Theorem 5.2.1.2. *For fixed $x \in \mathbb{R}$, let the process X satisfy the stochastic differential equation* (5.1)*. We assume that the coefficients μ and σ satisfy Condition 5.2.1.1. Moreover, let $(\widetilde{X}_t,\, t \geq 0)$ denote the solution to*

$$d\widetilde{X}_t = \mu(x)dt + \sigma(x)dB_t, \quad X_0 = x, \ t \geq 0. \tag{5.3}$$

Then we have

$$\left| \mathbb{E}_x\left[\sup_{0\leq s\leq t} X_s\right] - \mathbb{E}_x\left[\sup_{0\leq s\leq t} \widetilde{X}_s\right] \right| = O(t). \tag{5.4}$$

Proof. To begin with, let us note that

$$\mathbb{E}_x\left[\left|\sup_{0\leq s\leq t} X_s - \sup_{0\leq s\leq t} \widetilde{X}_s\right|\right] \leq \mathbb{E}_x\left[\sup_{0\leq s\leq t} |X_s - \widetilde{X}_s|\right]. \tag{5.5}$$

We want to estimate $\mathbb{E}_x\left[\sup_{0\leq s\leq t}|X_s - \widetilde{X}_s|^2\right]$. By means of Young's inequality, in a first step we obtain

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\leq s\leq t}|X_s - \widetilde{X}_s|^2\right] &\leq 2\mathbb{E}_x\left[\sup_{0\leq s\leq t}\left|\int_0^s (\mu(X_u) - \mu(x))du\right|^2\right] \\ &\quad + 2\mathbb{E}_x\left[\sup_{0\leq s\leq t}\left|\int_0^s \left(\sigma(X_u) - \sigma(x)\right)dB_u\right|^2\right].\end{aligned} \tag{5.6}$$

The second term on the right hand side can be estimated by means of Doob's inequality and the Lipschitz property of σ. Note that Condition 5.2.1.1 implies that μ and σ are Lipschitz continuous with uniform Lipschitz constant. We obtain

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\le s\le t}\left|\int_0^s\left(\sigma(X_u)-\sigma(x)\right)dB_u\right|^2\right] &\preceq \sup_{0\le s\le t}\mathbb{E}_x\left[\left(\int_0^s\left(\sigma(X_u)-\sigma(x)\right)dB_u\right)^2\right]\\ &\preceq t\sup_{0\le s\le t}\mathbb{E}_x\left[\left|X_s-x\right|^2\right].\end{aligned}\tag{5.7}$$

The symbol $\preceq$ means "less or equal but up to a positive constant". The first term on the right hand side of (5.6) can be estimated by means of Cauchy-Schwarz' inequality and by the Lipschitz property of μ. The resulting estimate is

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\le s\le t}\left|\int_0^s(\mu(X_u)-\mu(x))du\right|^2\right] &\le t\,\mathbb{E}_x\left[\sup_{0\le s\le t}\left|\int_0^s(\mu(X_u)-\mu(x))^2du\right|\right]\\ &\preceq t\,\mathbb{E}_x\left[\int_0^t(X_u-x)^2du\right]\\ &\le t^2\sup_{0\le s\le t}\mathbb{E}_x\left[(X_u-x)^2\right].\end{aligned}\tag{5.8}$$

Bringing the previous estimates together, one obtains the overall estimate

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\le s\le t}\left|X_s-\widetilde{X}_s\right|^2\right] &\le 2\mathbb{E}_x\left[\sup_{0\le s\le t}\left|\int_0^s(\mu(X_u)-\mu(x))du\right|^2\right]\\ &\quad+2\mathbb{E}_x\left[\sup_{0\le s\le t}\left|\int_0^s\left(\sigma(X_u)-\sigma(x)\right)dB_u\right|^2\right]\\ &\preceq t^2\sup_{0\le s\le t}\mathbb{E}_x\left[\left|X_s-x\right|^2\right]+t\sup_{0\le s\le t}\mathbb{E}_x\left[\left|X_s-x\right|^2\right]\\ &\preceq t^2\cdot t+t\cdot t\sim t^2.\end{aligned}\tag{5.9}$$

By Cauchy-Schwarz' inequality the estimates (5.5) and (5.9) give the assertion. □

We introduce the following notations:

$$H_t=\sup_{0\le s\le t}X_s\tag{5.10}$$

and

$$H_t^B=\sup_{0\le s\le t}B_s.\tag{5.11}$$

Furthermore, we say that a function $g:\mathbb{R}^d\to\mathbb{R}$ has *polynomial growth near infinity*, if there exists a vector of integers $(k_1,\dots,k_d)$ such that $|y_1|^{-k_1}\cdot\ldots\cdot|y_d|^{-k_d}g(y)$ is bounded above in a neighborhood of infinity. An immediate consequence of the Theorem 5.2.1.2 is stated in the following corollary.

Corollary 5.2.1.3. *Let the X satisfy the stochastic differential equation* (5.1). *We assume that the coefficients μ and σ satisfy Condition 5.2.1.1. Let $g \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ and suppose that g and both of its derivatives have polynomial growth near infinity, then*

$$\begin{aligned}\mathbb{E}_x[g(H_t)] &= g(x) + g'(x)\sigma(x)\mathbb{E}_0[H_t^B] + O(t) \\ &= g(x) + g'(x)\sigma(x)\sqrt{\frac{2}{\pi}}\sqrt{t} + O(t), \qquad (5.12)\end{aligned}$$

as $t \to 0$. Here, H and H^B are defined by (5.10) *and* (5.11), *respectively.*

Proof. The estimates in the proof of Theorem 5.2.1.2 show that

$$\Big|\mathbb{E}_x[H_t] - x - \sigma(x)\mathbb{E}_0[H_t^B]\Big| = O(t). \qquad (5.13)$$

Moreover, the following inequality can be proved along the same lines as the estimate (5.9) in the proof of Theorem 5.2.1.2:

$$\begin{aligned}\mathbb{E}_x\left[\left(\sup_{0\le s\le t} X_s - x\right)^2\right] &\le 2\,\mathbb{E}_x\left[\left(\sup_{0\le s\le t}\int_0^s \mu(X_u)du\right)^2\right] \\ &\quad + 2\,\mathbb{E}_x\left[\left(\sup_{0\le s\le t}\int_0^s \sigma(X_u)dB_u\right)^2\right] \\ &\preceq t^2 + t. \qquad (5.14)\end{aligned}$$

This inequality shows that

$$\mathbb{E}_x[(H_t - x)^2] = O(t). \qquad (5.15)$$

On the other hand, for $x, y \in \mathbb{R}$, we have

$$g(y) = g(x) + g'(x)(y-x) + \frac{1}{2}g''(\xi)(y-x)^2, \qquad (5.16)$$

where ξ is between x and y. Since g'' is continuous and has polynomial growth near infinity, we obtain the estimate

$$\begin{aligned}&\Big|\mathbb{E}_x[g(H_t)] - \Big\{g(x) + g'(x)\sigma(x)\mathbb{E}_0[H_t^B]\Big\}\Big| \\ &\preceq g'(x)\Big|\mathbb{E}_x[H_t] - x - \sigma(x)\mathbb{E}_0[H_t^B]\Big| + \frac{1}{2}\mathbb{E}_x\Big[|g''(H_t)|\cdot(H_t - x)^2\Big]. \qquad (5.17)\end{aligned}$$

The first term on the right hand side of (5.17) belongs to $O(t)$. This follows directly from (5.13). The second term on the right hand side of (5.17) can easily be estimated by means of equation (5.15) and Cauchy-Schwarz' inequality. By the fact that the function

g'' is continuous and has polynomial growth, it is equally easy to see that this term also belongs to $O(t)$. □

Let us continue our analysis. We state an auxiliary lemma, that will turn out to be crucial in the sequel.

Lemma 5.2.1.4. *Let X be a diffusion process that satisfies the stochastic differential equation*

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = 0, \; t \geq 0, \tag{5.18}$$

with coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ that satisfy Condition 5.2.1.1. Let H_t and H_t^B be defined as in (5.10) *and* (5.11)*, respectively. Moreover, define*

$$L_t = \inf_{0 \leq s \leq t} X_s \tag{5.19}$$

and

$$L_t^B = \inf_{0 \leq s \leq t} B_s. \tag{5.20}$$

Let $A_t^{(1)}, A_t^{(2)} \in \{H_t, L_t, X_t\}$ and let

$$A_t^{(i,B)} = \begin{cases} H_t^B, & \text{if } A_t^{(i)} = H_t, \\ L_t^B, & \text{if } A_t^{(i)} = L_t, \\ B_t, & \text{if } A_t^{(i)} = X_t, \end{cases} \tag{5.21}$$

for $i = 1, 2$. Then, for $m, n \in \mathbb{N}$,

$$\begin{aligned} &\mathbb{E}_x\Big[\left(A_t^{(1)} - x\right)^m \left(A_t^{(2)} - x\right)^n \Big] \\ &\qquad = \sigma^{(m+n)}(x)\mathbb{E}_0\Big[\left(A_t^{(1,B)}\right)^m \left(A_t^{(2,B)}\right)^n \Big] + O(t^{(m+n+1)/2}). \end{aligned} \tag{5.22}$$

Moreover, the refined formula

$$\begin{aligned} &\mathbb{E}_x\Big[\left(A_t^{(1)} - \mathbb{E}_x\left[A_t^{(1)}\right]\right)^m \left(A_t^{(2)} - \mathbb{E}_x\left[A_t^{(2)}\right]\right)^n \Big] \\ &\qquad = \sigma^{(m+n)}(x)\mathbb{E}_x\Big[\left(A_t^{(1,B)} - \mathbb{E}_x\Big[(A_t^{(1,B)}\Big]\right)^m \left(A_t^{(2,B)} - \mathbb{E}_x\Big[(A_t^{(2,B)}\Big]\right)^n \Big] \\ &\qquad\quad + \; O(t^{(m+n+1)/2}) \end{aligned} \tag{5.23}$$

holds.

Remark 5.2.1.5. Let us note that it is not clear whether the moments of (H_t^B, L_t^B) can be derived in explicit form. As we have already mentioned several times, this is possible for $\mathbb{E}_0[H_t^B L_t^B] = t(1 - 2\log 2)$. This result was proved by Rogers and Shepp, see [62].

However, higher order moments like $\mathbb{E}_0\Big[(H_t^B)^m(L_t^B)^n\Big],\ m,n\in\mathbb{N}$, can easily be approximated by simulations. Some moments are displayed in the following table. The values rely on a simulation of $5\cdot 10^5$ trajectories of $(B_t,\,0\le t\le 1)$ and the interval $[0,1]$ was split into 10^6 equidistant subintervals.

$\mathbb{E}[H_1^m\cdot L_1^n]$	m=1	m=2	m=3	m=4
n=1	-0.386	-0.342	-0.43	-0.685
n=2	0.343	0.227	0.226	0.304
n=3	-0.432	-0.227	-0.181	-0.202
n=4	0.691	0.304	0.201	0.185

Table 5.1: Simulated moments of the vector $(H_1,L_1)=\big(\sup_{0\le t\le 1}B_s,\,\inf_{0\le t\le 1}B_s\big)$.

Proof (of Lemma 5.2.1.4). Without loss of generality, let us assume that $x=0$. Simple arithmetics yield

$$\begin{aligned}
&\mathbb{E}_0\left[\left(A_t^{(1)}\right)^m\left(A_t^{(2)}\right)^n-\sigma^{(m+n)}(0)\left(A_t^{(1,B)}\right)^m\left(A_t^{(2,B)}\right)^n\right]\\
&=\mathbb{E}_0\left[\left(A_t^{(1)}\right)^m\left(A_t^{(2)}\right)^n-\sigma^{n}(0)\left(A_t^{(1)}\right)^m\left(A_t^{(2,B)}\right)^n\right]\\
&\qquad+\mathbb{E}_0\left[\sigma^n(0)\left(A_t^{(1)}\right)^m\left(A_t^{(2,B)}\right)^n-\sigma^{(m+n)}(0)\left(A_t^{(1,B)}\right)^m\left(A_t^{(2,B)}\right)^n\right]\\
&=\mathbb{E}_0\left[\left(A_t^{(1)}\right)^m\left\{\left(A_t^{(2)}\right)^n-\sigma^n(0)\left(A_t^{(2,B)}\right)^n\right\}\right]\\
&\qquad+\mathbb{E}_0\left[\left\{\left(A_t^{(1)}\right)^m-\sigma^m(0)\left(A_t^{(1,B)}\right)^m\right\}\sigma^n(0)\left(A_t^{(2,B)}\right)^n\right].
\end{aligned}\tag{5.24}$$

Clearly, the following two equations hold:

$$\left(A_t^{(1)}\right)^m-\sigma^m(0)\left(A_t^{(1,B)}\right)^m=\left(A_t^{(1)}-\sigma(0)A_t^{(1,B)}\right)\sum_{k=1}^{m-1}\left(A_t^{(2)}\right)^k\left(\sigma(0)A_t^{(2,B)}\right)^{m-k-1}\tag{5.25}$$

and

$$\left(A_t^{(2)}\right)^n-\sigma^n(0)\left(A_t^{(2,B)}\right)^n=\left(A_t^{(2)}-\sigma(0)A_t^{(2,B)}\right)\sum_{k=1}^{n-1}\left(A_t^{(2)}\right)^k\left(\sigma(0)A_t^{(2,B)}\right)^{n-k-1}.\tag{5.26}$$

Inequality (5.9) shows that

$$\mathbb{E}_x\left[\sup_{0\le s\le t}\left|X_s-\sigma(x)B_s\right|^2\right]\preceq t.\tag{5.27}$$

Particularly, the previous estimate (5.27) holds for $x = 0$. And since

$$\left|A_t^{(i)} - \sigma(0)A_t^{(i,B)}\right| \leq \sup_{0\leq s\leq t} |X_s - \sigma(0)B_s|, \tag{5.28}$$

for $i = 1, 2$, inequality (5.27) implies that

$$\sqrt{\mathbb{E}_0\left[\left(A_t^{(i)} - \sigma(0)A_t^{(i,B)}\right)^2\right]} = O(t), \tag{5.29}$$

for $i = 1, 2$. The result now follows directly by applying Cauchy-Schwarz' inequality to (5.24) repeatedly. The refinement follows easily by the fact that

$$\mathbb{E}_x\left[H_t\right] = x + \sigma(x)\mathbb{E}_0[H_t^B] + O(t) \tag{5.30}$$

and

$$\mathbb{E}_x\left[X_t\right] = x + O(t). \tag{5.31}$$

Consequently, both assertions in the above lemma are proved. $\square$

We are now going to study a very particular case. If the diffusion coefficient σ is a positive constant, we are able to derive a result that is stronger than the one of Theorem 5.2.1.2. This result is stated in the next theorem.

Theorem 5.2.1.6. *Let $x \in \mathbb{R}$ be fixed and let X be a diffusion that satisfies the stochastic differential equation*

$$dX_t = \mu(X_t)dt + \sigma\, dB_t, \quad X_0 = x,\ t \geq 0, \tag{5.32}$$

where B is the standard Brownian motion of $\mathbb{R}$ and $\sigma > 0$ is a constant. The drift coefficient $\mu : \mathbb{R} \to \mathbb{R}$ is supposed to satisfy Condition 5.2.1.1. Moreover, let $\widetilde{X}_t$ denote the solution to

$$d\widetilde{X}_t = \mu(x)dt + \sigma dB_t, \quad X_0 = x,\ t \geq 0. \tag{5.33}$$

We then have

$$\left|\mathbb{E}_x\left[\sup_{0\leq s\leq t} X_s\right] - \mathbb{E}_x\left[\sup_{0\leq s\leq t} \widetilde{X}_s\right]\right| = O(t^{3/2}). \tag{5.34}$$

Proof. The same proceeding as in the proof of Theorem 5.2.1.2 shows that

$$\begin{aligned}\mathbb{E}_x\left[\sup_{0\leq s\leq t} |X_s - \widetilde{X}_s|^2\right] &\leq 2\mathbb{E}_x\left[\sup_{0\leq s\leq t}\left|\int_0^s (\mu(X_u) - \mu(x))du\right|^2\right] \\ &\quad + 2\mathbb{E}_x\left[\sup_{0\leq s\leq t}\left|\int_0^s (\sigma - \sigma)\, dB_u\right|^2\right]\end{aligned}$$

$$\begin{aligned} &\preceq t^2 \sup_{0\le s\le t} \mathbb{E}_x\left[\left|X_s - x\right|^2\right] \\ &\preceq t^2 \cdot t = t^3. \end{aligned} \tag{5.35}$$

Especially, compare formula (5.9). Since

$$\left|\mathbb{E}_x\left[\sup_{0\le s\le t} X_s\right] - \mathbb{E}_x\left[\sup_{0\le s\le t} \widetilde{X}_s\right]\right| \le \mathbb{E}_x\left[\sup_{0\le s\le t} \left|X_s - \widetilde{X}_s\right|\right], \tag{5.36}$$

the result follows directly by applying Cauchy-Schwarz' inequality. □

By means of Theorem 5.2.1.6 we can infer another important result, which is stated in the next corollary.

Corollary 5.2.1.7. *Let $x \in \mathbb{R}$ be fixed and let X satisfy the stochastic differential equation* (5.1)*. We assume that the diffusion coefficient σ is a positive constant and that the drift coefficient $\mu : \mathbb{R} \to \mathbb{R}$ satisfies Condition 5.2.1.1. Let $\widetilde{X}$ be defined by* (5.33)*. If $g \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$ and if g and all of its derivatives have polynomial growth near infinity, then*

$$\mathbb{E}_x\left[g\left(H_t\right)\right] = \mathbb{E}_x[g(\widetilde{H}_t)] + O(t^{3/2}), \tag{5.37}$$

where $H_t = \sup_{0\le s\le t} X_s$ and $\widetilde{H}_t = \sup_{0\le s\le t} \widetilde{X}_s$.

Proof. For $x, y \in \mathbb{R}$ we have

$$g(y) = g(x) + g'(x)(y-x) + \frac{1}{2}g''(x)(y-x)^2 + \frac{1}{6}g'''(\xi)(y-x)^3, \tag{5.38}$$

where ξ is between x and y. Since g''' is continuous and has polynomial growth, we obtain the estimate

$$\begin{aligned} &\left|\mathbb{E}_x[g(H_t)] - \mathbb{E}_x[g(\widetilde{H}_t)]\right| \\ &\preceq g'(x)\left|\mathbb{E}_x[H_t - x] - \mathbb{E}_x[\widetilde{H}_t - x]\right| + \frac{1}{2}g''(x)\left|\mathbb{E}_0\left[H_t^2\right] - \mathbb{E}_0\left[\widetilde{H}_t^2\right]\right| \\ &\quad + \frac{1}{6}\mathbb{E}_x\left[\left|g'''(H_t)\right|(H_t - x)^3\right] + \frac{1}{6}\mathbb{E}_x\left[\left|g'''(\widetilde{H}_t)\right|(\widetilde{H}_t - x)^3\right]. \end{aligned} \tag{5.39}$$

By Theorem 5.2.1.6 we have

$$\left|\mathbb{E}_x[H_t - x] - \mathbb{E}_x[\widetilde{H}_t - x]\right| = O(t^{3/2}). \tag{5.40}$$

Moreover, from Lemma 5.2.1.4 we are able to infer that

$$\left|\mathbb{E}_x\left[(H_t - x)^2\right] - \mathbb{E}_x\left[(\widetilde{H}_t - x)^2\right]\right| = O(t^{3/2}). \tag{5.41}$$

The function g''' has polynomial growth and thus, by means of Cauchy-Schwarz' inequality and inequality (5.14), it is straightforward to show that the expressions

$$\frac{1}{6}\mathbb{E}_x\left[|g'''(H_t)|(H_t-x)^3\right] \quad \text{and} \quad \frac{1}{6}\mathbb{E}_x\left[|g'''(\widetilde{H}_t)|(\widetilde{H}_t-x)^3\right] \tag{5.42}$$

on the right hand side of (5.39) belong to $O(t^{3/2})$. This completes the proof of the corollary. □

The last corollary states that, for a sufficiently smooth function g and for a diffusion with constant diffusion coefficient $\sigma > 0$, the following relation holds

$$\mathbb{E}_x\left[g\left(H_t\right)\right] = \mathbb{E}_x\left[g\left(\widetilde{H}_t\right)\right] + O(t^{3/2}). \tag{5.43}$$

Here, $H_t = \sup_{0\le s\le t} X_s$, $\widetilde{H}_t = \sup_{0\le s\le t}\widetilde{X}_s$ and $\widetilde{X}$ denotes the Brownian approximation (5.33). Formula (5.43) shows that the coefficients belonging to $\sqrt{t}$ and to t in the expansion of $\mathbb{E}_x[g(H_t)]$ coincide with the respective terms in the expansion of $\mathbb{E}_x[g(\widetilde{H}_t)]$. Thus, if we were able to expand the moments of the running maximum $\widetilde{H}_t$ of a Brownian motion with drift, we would be able to derive the corresponding expansion for the running maximum H_t of the more general diffusion process X. The coefficient belonging to $\sqrt{t}$ was already determined in Corollary 5.2.1.3. It can be calculated by means of elementary properties of Brownian motion. Up to now, we have no tool to determine the coefficient belonging to t. This problem will be addressed in the next paragraph. But before, let us consider another particular example.

In subsequent chapters, we are sometimes going to consider the *Lamperti transform* of a diffusion process. Assume that X satisfies the stochastic differential equation (5.1). The Lamperti transform Y of X is formally defined by

$$Y_t = \int^{X_t} \frac{1}{\sigma(u)}du, \quad t \ge 0, \tag{5.44}$$

where any primitive of $1/\sigma(\cdot)$ may be selected. We have the following result, which is an immediate consequence of Theorem 5.2.1.6.

Corollary 5.2.1.8. *Let $x \in \mathbb{R}$ be fixed and consider a diffusion process X given by the stochastic differential equation* (5.1)*. We assume that $\sigma : \mathbb{R} \to \mathbb{R}_+$ is elliptic and at least once continuously differentiable. Let F be a primitive of $1/\sigma(\cdot)$. We define $Y = F(X)$ and $\xi = F(x)$ and we assume that the drift coefficient of the Lamperti transform, which is given by*

$$y \longmapsto \frac{\mu(F^{-1}(y))}{\sigma(F^{-1}(y))} - \frac{1}{2}\sigma'(F^{-1}(y)), \quad y \in \mathbb{R}, \tag{5.45}$$

satisfies Condition 5.2.1.1. If $\widetilde{Y}$ is a solution to

$$d\widetilde{Y}_t = \left(\frac{\mu(F^{-1}(\xi))}{\sigma(F^{-1}(\xi))} - \frac{1}{2}\sigma'(F^{-1}(\xi)) \right) dt + dB_t, \quad \widetilde{Y}_0 = \xi, \ t \geq 0, \tag{5.46}$$

then

$$\left| \mathbb{E}_x \left[\sup_{0\leq s\leq t} Y_s \right] - \mathbb{E}_x \left[\sup_{0\leq s\leq t} \widetilde{Y}_s \right] \right| = O(t^{3/2}). \tag{5.47}$$

Proof. The process Y starts in ξ and, by Itô's formula, it has diffusion coefficient 1. This follows because Y satisfies the stochastic differential equation

$$\begin{aligned} dY_t &= \frac{dX_t}{\sigma(X_t)} - \frac{1}{2}\frac{\sigma'(X_t)}{\sigma(X_t)^2} d\langle X \rangle_t \\ &= \left(\frac{\mu(X_t)}{\sigma(X_t)} - \frac{1}{2}\sigma'(X_t) \right) dt + dB_t \\ &= \left(\frac{\mu(F^{-1}(Y_t))}{\sigma(F^{-1}(Y_t))} - \frac{1}{2}\sigma'(F^{-1}(Y_t)) \right) dt + dB_t, \ t \geq 0. \end{aligned} \tag{5.48}$$

Due to the assumptions we made, the result now follows directly from Theorem 5.2.1.6. □

5.2.2 Expansions for the maximum of a Brownian motion with drift

Throughout this paragraph $(X_t,\, t \geq 0)$ denotes a Brownian motion with drift, that is $X_t = \mu t + \sigma B_t$ with $\mu \in \mathbb{R}$ and $\sigma > 0$. Moreover, we assume that $X_0 = x$ for a fixed value $x \in \mathbb{R}$. For a sufficiently smooth function g, we will now try to calculate an expansion of $\mathbb{E}_x[g(H_t)]$, where $H_t = \sup_{0\leq s\leq t} X_s$. As we have already mentioned in Section 3.4, in the Brownian case, the joint density of the maximum H_t and the terminal value X_t is given by

$$f_{(\mu,\sigma)}(t,x,h,y) = \frac{2(2h-x-y)}{\sqrt{2\pi t^3}\sigma^3} \exp\left(\frac{\mu}{\sigma^2}(y-x) - \frac{(2h-x-y)^2}{2t\sigma^2} - \frac{\mu^2}{2\sigma^2}t \right). \tag{5.49}$$

An integration of this density with respect to y yields the density of H_t, which is given by

$$\begin{aligned} f_{(\mu,\sigma)}(t,x,h) &= \int_{-\infty}^{h} f_{(\mu,\sigma)}(t,x,h,y)\, dy \\ &= \frac{\sqrt{\frac{2}{\pi}} \exp\left(-\frac{(h-x-t\mu)^2}{2t\sigma^2}\right) t\sigma^2 - \mu\sigma\sqrt{t^3} \left(1 - \mathrm{Erf}\left[\frac{(h-x+t\mu)}{\sqrt{2t}\sigma}\right]\right)}{\sqrt{t^3\sigma^6}}. \end{aligned} \tag{5.50}$$

The expression Erf denotes the error-function

$$z \longmapsto \operatorname{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp\left(-s^2\right) ds, \quad z \in \mathbb{R}. \tag{5.51}$$

Note that $\operatorname{Erf}(\cdot)$ is defined in such a way that integration starts in 0. Hence, we particularly have $\operatorname{Erf}(z) = O(z)$. For our purposes it is more convenient to work with the error-function $\operatorname{Erf}(\cdot)$ than with the cumulative density function $\Phi(\cdot)$ of a standard normal random variable, since this simplifies our notations. Tedious but routine calculations show that

$$\frac{\partial}{\partial t} f_{(\mu,\sigma)}(t,x,h) = -\mu \frac{\partial}{\partial h} f_{(\mu,\sigma)}(t,x,h) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial h^2} f_{(\mu,\sigma)}(t,x,h). \tag{5.52}$$

More precisely, we have

$$\frac{\partial}{\partial t} f_{(\mu,\sigma)}(t,x,h) = \frac{e^{-\frac{(-h+x+t\mu)^2}{2t\sigma^2}}\left((h-x)(h-x-t\mu) - t\sigma^2\right)}{\sqrt{2\pi}\sqrt{t^5\sigma^6}}, \tag{5.53}$$

whereas the first and second derivates with respect to h are given by

$$\begin{aligned} &\frac{\partial}{\partial h} f_{(\mu,\sigma)}(t,x,h) \\ &= -\frac{\sqrt{\frac{2}{\pi}} e^{-\frac{(-h+x+t\mu)^2}{2t\sigma^2}}(h-x-2t\mu)\sigma + 2\mu^2\sqrt{t^3}\left(1 - \operatorname{Erf}\left[\frac{(h-x+t\mu)}{\sqrt{2t}\sigma}\right]\right)}{\sigma\sqrt{t^3\sigma^6}} \end{aligned} \tag{5.54}$$

and

$$\begin{aligned} \frac{\partial^2}{\partial h^2} f_{(\mu,\sigma)}(t,x,h) = -e^{-\frac{(-h+x+t\mu)^2}{2t\sigma^2}} &\left(2\frac{(h-x)^2 - 3(h-x)t\mu + 4t^2\mu^2 - t\sigma^2}{\sigma^5\sqrt{2\pi}\sqrt{t^5}}\right) \\ &- 4\frac{\mu^3}{\sigma^6}\left(1 - \operatorname{Erf}\left[\frac{h-x+t\mu}{\sqrt{2}\sqrt{t}\sigma}\right]\right). \end{aligned} \tag{5.55}$$

If $g : \mathbb{R} \to \mathbb{R}$ is a sufficiently smooth function that does not grow too fast, formula (5.52) enables us to perform the following calculation

$$\begin{aligned} &\frac{\partial}{\partial t}\mathbb{E}_x[g(H_t)] \\ &= \int_x^\infty g(h)\frac{\partial}{\partial t} f_{(\mu,\sigma)}(t,x,h)dh \\ &= -\mu\int_x^\infty g(h)\frac{\partial}{\partial h} f_{(\mu,\sigma)}(t,x,h)dh + \frac{1}{2}\sigma^2\int_x^\infty g(h)\frac{\partial^2}{\partial h^2} f_{(\mu,\sigma)}(t,x,h)dh \\ &= -\mu\Big[g(h)f_{(\mu,\sigma)}(t,x,h)\Big]_{h=x}^{h=\infty} + \mu\int_x^\infty \frac{\partial}{\partial h} g(h) f_{(\mu,\sigma)}(t,x,h)dh \end{aligned}$$

$$+\frac{1}{2}\sigma^2\left[g(h)\frac{\partial}{\partial h}f_{(\mu,\sigma)}(t,x,h)\right]_{h=x}^{h=\infty}-\frac{1}{2}\sigma^2\left[\frac{\partial}{\partial h}g(h)f_{(\mu,\sigma)}(t,x,h)\right]_{h=x}^{h=\infty}$$
$$+\frac{1}{2}\sigma^2\int_x^\infty\frac{\partial^2}{\partial h^2}g(h)f_{(\mu,\sigma)}(t,x,h)dh. \tag{5.56}$$

Note that, on mild regularity assumptions, one is allowed to reverse the order of differentiation and integration in the first line of (5.56). A criterion is given in Chapter 11, *Satz 2* in the book of Forster [27]. Distinctly, for the special Brownian density (5.49), the assumptions of Forster's theorem are satisfied.

The function $f_{(\mu,\sigma)}(t,x,h)$ is a density and it satisfies

$$\mathbb{P}_x[H_t\geq h]=\int_h^\infty f_{(\mu,\sigma)}(t,x,a)da=\begin{cases}\longrightarrow 0, \text{ if } x<h,\\ \longrightarrow 1, \text{ if } x=h,\end{cases} \tag{5.57}$$

as $t\to 0$. From distribution theory, it follows that

$$\mu\int_x^\infty\frac{\partial}{\partial h}g(h)f_{(\mu,\sigma)}(t,x,h)dh\longrightarrow\mu\frac{\partial}{\partial x}g(x), \tag{5.58}$$

as $t\to 0$ and if g' is continuous in x. By the same reasoning it follows that

$$\frac{1}{2}\sigma^2\int_x^\infty\frac{\partial^2}{\partial h^2}g(h)f_{(\mu,\sigma)}(t,x,h)dh\longrightarrow\frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}g(x), \tag{5.59}$$

as $t\to 0$ and if g'' is continuous in x. But what happens to the remaining terms on the right hand side of (5.56)? First, it is obvious from formulae (5.50) and (5.54) that, for $x\in\mathbb{R}$ and $t>0$,

$$\lim_{h\to\infty}\frac{\partial}{\partial h}f_{(\mu,\sigma)}(t,x,h)=\lim_{h\to\infty}f_{(\mu,\sigma)}(t,x,h)=0. \tag{5.60}$$

Moreover,

$$f_{(\mu,\sigma)}(t,x,x)=\frac{e^{-\frac{t\mu}{2\sigma^2}}\sqrt{\frac{2}{\pi}}\sigma-\sqrt{t}\mu\left(1-\text{Erf}\left[\frac{\sqrt{t}\mu}{\sqrt{2}\sigma}\right]\right)}{\sqrt{t}\sigma^2}, \tag{5.61}$$

which implies

$$\mu\, g(x)f_{(\mu,\sigma)}(t,x,x)=\mu\, g(x)\sqrt{\frac{2}{\pi\sigma^2 t}}e^{-\frac{t\mu}{2\sigma^2}}-g(x)\frac{\mu^2}{\sigma^2}\left(1-\text{Erf}\left[\frac{\sqrt{t}\mu}{\sqrt{2}\sigma}\right]\right). \tag{5.62}$$

The first derivative of f with respect to h satisfies

$$\frac{\partial}{\partial h}f_{(\mu,\sigma)}(t,x,h)\Big|_{h=x}=2\frac{\sqrt{\frac{2}{\pi}}e^{-\frac{t\mu}{2\sigma^2}}t\mu\sigma\ -\mu^2\sqrt{t^3}\left(1-\text{Erf}\left[\frac{t\mu}{\sqrt{2t\sigma}}\right]\right)}{\sigma\sqrt{t^3\sigma^6}}, \tag{5.63}$$

which yields

$$\begin{aligned}
&\frac{1}{2}\sigma^2 g(x)\frac{\partial}{\partial h}f_{(\mu,\sigma)}(t,x,h)\Big|_{h=x} \\
&\quad = g(x)\sqrt{\frac{2}{\pi\sigma^2 t}}\,\mu e^{-\frac{t\mu}{2\sigma^2}} - g(x)\frac{\mu^2}{\sigma^2}\left(1-\mathrm{Erf}\left[\frac{t\mu}{\sqrt{2t}\sigma}\right]\right). \qquad (5.64)
\end{aligned}$$

Obviously (5.62) and (5.64) are equal and thus the corresponding expressions cancel each other in formula (5.56).

Lastly, from (5.61) we have

$$\frac{1}{2}\sigma^2 g'(x)f_{(\mu,\sigma)}(t,x,x) = \frac{1}{2}\sigma^2 g'(x)\sqrt{\frac{2}{\pi\sigma^2 t}}e^{-\frac{t\mu}{2\sigma^2}} - \frac{1}{2}g'(x)\mu\left(1-\mathrm{Erf}\left[\frac{\sqrt{t}\mu}{\sqrt{2}\sigma}\right]\right). \qquad (5.65)$$

This gives the intermediate result

$$\begin{aligned}
\frac{\partial}{\partial t}\mathbb{E}_x[g(H_t)] &= \frac{1}{2}\sigma^2 g'(x)\sqrt{\frac{2}{\pi\sigma^2 t}}e^{-\frac{t\mu}{2\sigma^2}} - \frac{1}{2}g'(x)\mu\left(1-\mathrm{Erf}\left[\frac{\sqrt{t}\mu}{\sqrt{2}\sigma}\right]\right) \\
&\quad + \mu\int_x^\infty \frac{\partial}{\partial h}g(h)f_{(\mu,\sigma)}(t,x,h)dh \\
&\quad + \frac{1}{2}\sigma^2\int_x^\infty \frac{\partial^2}{\partial h^2}g(h)f_{(\mu,\sigma)}(t,x,h)dh. \qquad (5.66)
\end{aligned}$$

By letting $t \to 0$ we find the following result.

Theorem 5.2.2.1. *Let $g : \mathbb{R} \to \mathbb{R}$ a three times continously differentiable function such that g and all of its derivatives have polynomial growth near infinity. Moreover, let H denote the running maximum process of X, where $(X_t = \mu t + \sigma B_t,\ t \geq 0)$ is a Brownian motion with drift. Finally, let $x \in \mathbb{R}$. Then we have*

$$\begin{aligned}
\mathbb{E}_x[g(H_t)] &= g(x) + \sqrt{t}g'(x)\frac{2\sigma}{\sqrt{2\pi}} + t\left(\frac{1}{2}\mu\frac{\partial}{\partial x}g(x) + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}g(x)\right) \\
&\quad + O(t^{3/2}). \qquad (5.67)
\end{aligned}$$

Proof. First, note that the polynomial growth condition ensures the existence of $\mathbb{E}_x[g(H_t)]$, $\mathbb{E}_x[g'(H_t)]$ and $\mathbb{E}_x[g''(H_t)]$. Let us define the second order expansion

$$T_2(\mathbb{E}_x[g(H_t)]) := g(x) + 2\sqrt{t}\left(g'(x)\frac{\sigma}{\sqrt{2\pi}}\right) + t\left(\frac{1}{2}\mu\frac{\partial}{\partial x}g(x) + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}g(x)\right). \qquad (5.68)$$

By differentiating this expression with respect to t and comparing it with (5.66) one finds that

$$\begin{aligned}
&\frac{\partial}{\partial t} T_2(\mathbb{E}_x[g(H_t)]) - \frac{\partial}{\partial t}\mathbb{E}_x[g(H_t)] \\
&= \frac{1}{2}\sigma^2 g'(x)\sqrt{\frac{2}{\pi\sigma^2 t}}\left(1 - e^{-\frac{t\mu}{2\sigma^2}}\right) + \mu g'(x) - \mu \int_x^\infty g'(h) f_{(\mu,\sigma)}(t,x,h)dh \\
&\qquad + \frac{1}{2}\sigma^2 g''(x) - \frac{1}{2}\sigma^2 \int_x^\infty g''(h) f_{(\mu,\sigma)}(t,x,h)dh - \frac{1}{2}\mu g'(x)\mathrm{Erf}\left[\frac{\sqrt{t}\mu}{\sqrt{2}\sigma}\right].
\end{aligned} \tag{5.69}$$

First, it is obvious that

$$\left(1 - e^{-\frac{t\mu}{2\sigma^2}}\right) = O(t), \tag{5.70}$$

or equivalently

$$\frac{1}{\sqrt{t}}\left(1 - e^{-\frac{t\mu}{2\sigma^2}}\right) = O(\sqrt{t}). \tag{5.71}$$

Furthermore, from Theorem 5.2.1.2 we derive that

$$\int_x^\infty g'(h) f_{(\mu,\sigma)}(t,x,h)dh = \mathbb{E}_x[g'(H_t)] = g'(x) + O(\sqrt{t}) \tag{5.72}$$

and

$$\int_x^\infty g''(h) f_{(\mu,\sigma)}(t,x,h)dh = \mathbb{E}_x[g''(H_t)] = g''(x) + O(\sqrt{t}). \tag{5.73}$$

Finally let us state that

$$\mathrm{Erf}\left[\frac{\sqrt{t}\mu}{\sqrt{2}\sigma}\right] = O(\sqrt{t}). \tag{5.74}$$

This follows directly from the definition of $\mathrm{Erf}(\cdot)$ in (5.51). Altogether, we obtain the following estimate

$$\frac{\partial}{\partial t} T_2(\mathbb{E}_x[g(H_t)]) - \frac{\partial}{\partial t}\mathbb{E}_x[g(H_t)] = O(\sqrt{t}), \tag{5.75}$$

which yields the desired result. □

In order to close this paragraph, let us state three more or less obvious observations.

Remark 5.2.2.2. The proof of the previous theorem shows that the remainder term $O(t^{3/2})$ depends on the derivatives g', g'' and g''' of g. This contrasts with the situation of an ordinary expansion like the Taylor expansion, where the remainder term only depends on the highest order derivative.

Remark 5.2.2.3. A higher order expansion of Brownian motion with drift can be obtained in a similar way by differentiating the density (5.50) again with respect to t. The result will not be displayed here, since it does not provide fundamental insights.

Remark 5.2.2.4. For $g = id$ we have the particular expansion

$$\mathbb{E}_x\left[\sup_{0\leq s\leq t}(\mu s+\sigma B_s)\right] = x+\sqrt{t}\frac{2\sigma}{\sqrt{2\pi}}+\frac{1}{2}\mu t+O(t^{3/2}). \tag{5.76}$$

This result can be explained in the following way. The Brownian motion $(B_s,\, 0 \leq s \leq t)$ takes its maximum at $s = \frac{1}{2}t$ on average. In other words, the random time $\tau_t = \inf\{u \in [0,t] \,|\, B_u = \sup_{0\leq s\leq t} B_s\}$ satisfies $\mathbb{E}_x[\tau_t] = \frac{1}{2}t$. Note that τ_t is almost surely unique. For further details, see page 102 in the book of Karatzas and Shreve [43]. A lax interpretation of the law of the iterated logarithm suggests that, for small t, standard Brownian motion moves to its maximum and away from its maximum at a rate of $\sqrt{t}$. We were considering a Brownian motion with drift X, defined by $X_s = \mu s + \sigma B_s$, $0 \leq s \leq t$. If μ is positive, the decrease of the Brownian part away from its maximum outweighs the increase in the linear part. And for μ negative, the movement of the Brownian part to its maximum outweighs the decrease of the linear part. Our heuristic suggests that the process $(X_s,\, 0 \leq s \leq t)$ also takes its maximum value at about time $s = \frac{1}{2}t$ on average. This gives an intuitive explanation for the particular form of the second order term $\frac{1}{2}\mu t$ in formula (5.76).

5.2.3 Expansions for the maximum of a diffusion process

In this section we focus again on diffusion processes more general than Brownian motion. The aim is to combine the results of Section 5.2.1 and Section 5.2.2 in order to find an expansion of $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$. Here, X denotes a diffusion defined by the stochastic differential equation (5.1) and as usual the processes H and L are defined by $H_t = \sup_{0\leq s\leq t} X_s$ and $L_t = \inf_{0\leq s\leq t} X_s$. A first step towards our goal is the next theorem.

Theorem 5.2.3.1. *Let $x \in \mathbb{R}$ and let X satisfy the stochastic differential equation* (5.1)*. We assume that the diffusion coefficient σ is a positive constant and that the drift coefficient $\mu : \mathbb{R} \to \mathbb{R}$ satisfies Condition 5.2.1.1. If $g \in \mathcal{C}^3(\mathbb{R},\mathbb{R})$ and if g and all of its derivatives have polynomial growth near infinity, then the process* $(H_t = \sup_{0\leq s\leq t} X_s,\, t \geq 0)$ *satisfies*

$$\mathbb{E}_x\left[g\left(H_t\right)\right] = g(x)+g'(x)\sigma\sqrt{\frac{2}{\pi}}\sqrt{t}+g'(x)\frac{1}{2}\mu(x)t+g''(x)\frac{1}{2}\sigma^2 t+O(t^{3/2}). \tag{5.77}$$

Proof. Combining Corollary 5.2.1.7 and Theorem 5.2.2.1, one obtains the result. □

From the previous theorem we infer the following corollary concerning the Lamperti transform of a diffusion.

Corollary 5.2.3.2. *For $x \in \mathbb{R}$ fixed, let X be a solution to the stochastic differential equation* (5.1)*. Let $\sigma : \mathbb{R} \to \mathbb{R}_+$ be elliptic and once continuously differentiable, and let F be any primitive of $1/\sigma(\cdot)$. Let the process $Y = F(X)$ denote the Lamperti transform of X as defined in* (5.44)*. We set $\xi = F(X_0)$ and we assume that the assumptions of Corollary 5.2.1.8 are satisfied. Let $g : \mathbb{R} \to \mathbb{R}$ be a three times continuously differentiable function such that g and all of its derivatives have polynomial growth near infinity. Then the process*

$$H_t^Y = \sup_{0 \le s \le t} Y_t \tag{5.78}$$

has the following expansion with respect to $\sqrt{t}$:

$$\begin{aligned}\mathbb{E}_x[g(H_t^Y)] = g(\xi) + g'(\xi)\sqrt{\frac{2}{\pi}}\sqrt{t} + g'(\xi)\frac{1}{2}\Big(\frac{\mu(x)}{\sigma(x)} - \frac{1}{2}\sigma'(x)\Big)t + g''(\xi)\frac{1}{2}t \\ + O(t^{3/2}).\end{aligned} \tag{5.79}$$

Proof. Recall the specific form of the stochastic differential equation for Y

$$dY_t = \left(\frac{\mu(F^{-1}(Y_t))}{\sigma(F^{-1}(Y_t))} - \frac{1}{2}\sigma'(F^{-1}(Y_t))\right) dt + dB_t, \quad Y_0 = \xi, \ t \ge 0. \tag{5.80}$$

Given Theorem 5.2.2.1, we obtain the proof as an immediate consequence of Corollary 5.2.1.8. $\square$

Let $g : \mathbb{R}^d \to \mathbb{R}$ and let $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ be a multi-index. We set $|\alpha| = \alpha_1 + \ldots + \alpha_d$. Whenever it is convenient, we will write $g_\alpha(x)$ instead of

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} g(x). \tag{5.81}$$

Before we proceed, let us state an auxiliary result that is an immediate consequence of Itô's formula.

Lemma 5.2.3.3. *For $x \in \mathbb{R}$ fixed, let X be a diffusion defined by the stochastic differential equation* (5.1)*. Additionally, let μ and σ satisfy Condition 5.2.1.1. If $g : \mathbb{R}^3 \to \mathbb{R}$ belongs to the space $\mathcal{C}^{3,3,4}(\mathbb{R}^3, \mathbb{R})$ and if g and all of its partial derivatives have polynomial growth near infinity, then*

$$\begin{aligned}&\mathbb{E}_x[g(H_t, L_t, X_t)] \\ &= g(x,x,x) + g_{1,0,0}(x,x,x)\mathbb{E}_x\left[\int_0^t dH_s\right] + g_{2,0,0}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dH_v dH_s\right] \\ &\quad + g_{1,1,0}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dL_v dH_s\right] + g_{1,0,1}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dX_v dH_s\right] \\ &\quad + g_{0,1,0}(x,x,x)\mathbb{E}_x\left[\int_0^t dL_s\right] + g_{1,1,0}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dH_v dL_s\right]\end{aligned}$$

$$
\begin{aligned}
&+ g_{0,2,0}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dL_v dL_s\right] + g_{0,1,1}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dX_v dL_s\right]\\
&+ g_{0,0,1}(x,x,x)\mathbb{E}_x\left[\int_0^t dX_s\right] + g_{1,0,1}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dH_v dX_s\right]\\
&+ g_{0,1,1}(x,x,x)\mathbb{E}_x\left[\int_0^t\int_0^s dL_v dX_s\right] + g_{0,0,2}(x,x,x)\mathbb{E}_x\left[\int_0^t d\langle X\rangle_s\right]\\
&+ O(t^{3/2}),
\end{aligned}
\tag{5.82}
$$

where $H_t = \sup_{0\le s\le t} X_s$ and $L_t = \inf_{0\le s\le t} X_s$.

Proof. The proof requires a lot of tedious calculations. It can be found in Appendix 10.1. □

Remark 5.2.3.4. Note that a slight modification of the proof for Lemma 5.2.3.3 shows that the assumption $g \in \mathcal{C}^{3,3,3}$ is sufficient to state the result. However, the estimates are even more cumbersome in this case. We omit further details.

We are now able to state the main result of this section.

Theorem 5.2.3.5. *Let X be a diffusion that satisfies the stochastic differential equation*

$$dX_t = \mu(X_t)dt + \sigma\, dB_t, \quad X_0 = x,\ t \ge 0, \tag{5.83}$$

where B is the standard Brownian motion of $\mathbb{R}$ and $\sigma > 0$ is a constant. Let the drift coefficient μ satisfy Condition 5.2.1.1. If $g : \mathbb{R}^3 \to \mathbb{R}$ is a function that satisfies the assumptions of Lemma 5.2.3.3, then

$$
\begin{aligned}
&\mathbb{E}_x[g(H_t, L_t, X_t)]\\
&\quad = g(x,x,x)\\
&\qquad + \ g_{1,0,0}(x,x,x)\sigma\sqrt{\frac{2}{\pi}}\sqrt{t} + g_{1,0,0}(x,x,x)\frac{1}{2}\mu(x)t + g_{2,0,0}(x,x,x)\frac{1}{2}\sigma^2 t\\
&\qquad - \ g_{0,1,0}(x,x,x)\sigma\sqrt{\frac{2}{\pi}}\sqrt{t} + g_{0,1,0}(x,x,x)\frac{1}{2}\mu(x)t + g_{0,2,0}(x,x,x)\frac{1}{2}\sigma^2 t\\
&\qquad + \ (1-2\log 2)g_{1,1,0}(x,x,x)\sigma^2 t\\
&\qquad + \ \frac{1}{2}g_{1,0,1}(x,x,x)\sigma^2 t + \frac{1}{2}g_{0,1,1}(x,x,x)\sigma^2 t\\
&\qquad + \ g_{0,0,1}(x,x,x)\mu(x)t + \frac{1}{2}g_{0,0,2}(x,x,x)\sigma^2 t\\
&\qquad + \ O(t^{3/2}),
\end{aligned}
\tag{5.84}
$$

where $H_t = \sup_{0\le s\le t} X_s$ and $L_t = \inf_{0\le s\le t} X_s$.

Proof. The proof is straightforward, but requires very tedious calculations. Therefore, it was moved to Appendix 10.1. □

As an immediate consequence of the previous Theorem 5.2.3.5, one obtains the following corollary.

Corollary 5.2.3.6. *Let $x \in \mathbb{R}$ be fixed and let X be a solution to the stochastic differential equation* (5.1)*, where we assume that the diffusion coefficient $\sigma : \mathbb{R} \to \mathbb{R}_+$ is elliptic and at least once continuously differentiable. Let Y denote the Lamperti transform of X and let H^Y denote its running maximum as defined in* (5.44) *and* (5.78)*, respectively. Moreover, we set*

$$L_t^Y = \inf_{0 \le s \le t} Y_s. \tag{5.85}$$

We assume that the drift coefficient of the Lamperti transform, which is given by $(\mu/\sigma - \frac{1}{2}\sigma') \circ F^{-1}$, satisfies Condition 5.2.1.1. Let $g : \mathbb{R}^3 \to \mathbb{R}$ be a function that satisfies the assumptions of Lemma 5.2.3.3, then for $\xi = F(X_0) = F(x)$ we have

$$\begin{aligned}
&\mathbb{E}_x[g(H_t^Y, L_t^Y, Y_t)] \\
&\quad = g(\xi,\xi,\xi) \\
&\quad + g_{1,0,0}(\xi,\xi,\xi)\sqrt{\frac{2}{\pi}}\sqrt{t} + g_{1,0,0}(\xi,\xi,\xi)\frac{1}{2}\left(\frac{\mu(x)}{\sigma(x)} - \frac{1}{2}\sigma'(x)\right)t + g_{2,0,0}(\xi,\xi,\xi)\frac{1}{2}t \\
&\quad - g_{0,1,0}(\xi,\xi,\xi)\sqrt{\frac{2}{\pi}}\sqrt{t} + g_{0,1,0}(\xi,\xi,\xi)\frac{1}{2}\left(\frac{\mu(x)}{\sigma(x)} - \frac{1}{2}\sigma'(x)\right)t + g_{0,2,0}(\xi,\xi,\xi)\frac{1}{2}t \\
&\quad + (1 - 2\log 2)g_{1,1,0}(\xi,\xi,\xi)t \\
&\quad + \frac{1}{2}g_{1,0,1}(\xi,\xi,\xi)t + \frac{1}{2}g_{0,1,1}(\xi,\xi,\xi)t \\
&\quad + g_{0,0,1}(\xi,\xi,\xi)\left(\frac{\mu(x)}{\sigma(x)} - \frac{1}{2}\sigma'(x)\right)t + \frac{1}{2}g_{0,0,2}(\xi,\xi,\xi)t \\
&\quad + O(t^{3/2}).
\end{aligned} \tag{5.86}$$

Proof. Again, we stress that Y satisfies the stochastic differential equation

$$dY_t = \left(\frac{\mu(F^{-1}(Y_t))}{\sigma(F^{-1}(Y_t))} - \frac{1}{2}\sigma'(F^{-1}(Y_t))\right)dt + dB_t, \quad Y_0 = \xi,\ t \ge 0. \tag{5.87}$$

The result now follows from Lemma 5.2.3.3 in combination with the moments we calculated in the proof of Theorem 5.2.3.5. ☐

Remark 5.2.3.7. In the upcoming chapter on small-Δ-optimality, we will consider martingale estimating functions $g_{\Delta,\theta}(x,h,l,y)$ which depend on the time variable $t = \Delta$ as well as on the state variables (x,h,l,y). So far, this particular case has been neglected in our analysis. But we will see that, if we impose sensible assumptions on the function $g_{\Delta,\theta}$, we are not limited to the results developed so far.

6 Small-Delta-Optimality

6.1 Introduction & Motivation

In Chapter 4 we considered martingale estimating functions for diffusion processes that satisfy the stochastic differential equation

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = x,\ t \geq 0, \tag{6.1}$$

with a real valued parameter $\theta \in \Theta \subset \mathbb{R}$. Suppose $g : \mathbb{R} \to \mathbb{R}$ is a $(2n+1)$-times continuously differentiable function. On mild regularity assumptions for the process X, the following expansion of the expectation holds

$$\mathbb{E}_{x,\theta}[g(X_t)] = \sum_{j=0}^{n} \frac{t^j}{j!}\mathcal{A}_\theta^j g(x) + o(t^n). \tag{6.2}$$

Here $\mathcal{A}_\theta$ denotes the infinitesimal generator of the process (6.1). One immediately obtains the approximations

$$\mathbb{E}_{x,\theta}[X_\Delta] = F(\Delta, x;\theta) = x + \Delta\mu(x;\theta) + o(\Delta) \tag{6.3}$$

and

$$\mathrm{Var}_{x,\theta}[X_\Delta] = \phi(\Delta, x;\theta) = \Delta\sigma^2(x;\theta) + o(\Delta). \tag{6.4}$$

The optimal linear ordinary estimating function, constructed from the sample vector $(X_{t_1}, \dots, X_{t_n})$, is given by

$$\sum_{i=1}^{n} \frac{\partial_\theta F(\Delta_i, X_{t_{i-1}};\theta)}{\phi(\Delta_i, X_{t_{i-1}};\theta)}[X_{t_i} - F(\Delta_i, X_{t_{i-1}};\theta)], \tag{6.5}$$

see the results of section 4.3.1. The calculation of the derivative $\partial_\theta F(\Delta_i, X_{t_{i-1}};\theta)$ is a non-trivial numerical problem, which we wish to avoid. For more details about this topic, see Pedersen [52]. By the above approximations of the expectation and of the variance, for small values of Δ, a reasonable approximate estimating function in the linear case is given by

$$\sum_{i=1}^{n} \frac{\partial_\theta \mu(X_{t_{i-1}};\theta)}{\sigma^2(X_{t_{i-1}};\theta)}[X_{t_i} - F(\Delta_i, X_{t_{i-1}};\theta)]. \tag{6.6}$$

In the quadratic case, we derived the approximate score function

$$\sum_{i=1}^{n} \frac{\partial_\theta F(\Delta_i, X_{t_{i-1}};\theta)}{\phi(\Delta_i, X_{t_{i-1}};\theta)}[X_{t_i} - F(\Delta_i, X_{t_{i-1}};\theta)] \\ + \sum_{i=1}^{n} \frac{\partial_\theta \phi(\Delta_i, X_{t_{i-1}};\theta)}{2\phi^2(\Delta_i, X_{t_{i-1}};\theta)}[(X_{t_i} - F(\Delta_i, X_{t_{i-1}};\theta))^2 - \phi(\Delta_i, X_{t_{i-1}};\theta)] \tag{6.7}$$

from a Brownian transition density with drift parameter $F(\Delta, x;\theta)$ and diffusion coefficient $\phi(\Delta, x;\theta)$. See formulae (4.13) and (4.14) in the introduction of Chapter 4. We see that the non-trivial derivatives $\partial_\theta F(\Delta_i, X_{t_{i-1}};\theta)$ and $\partial_\theta \phi(\Delta_i, X_{t_{i-1}};\theta)$ are involved. For small Δ, it seems reasonable to, once again, approximate (6.7) by a less complicated expression. A canonical choice is to consider the following martingale estimating function

$$\sum_{i=1}^{n} \frac{\partial_\theta \mu(X_{t_{i-1}};\theta)}{\sigma^2(X_{t_{i-1}};\theta)}[X_{t_i} - F(\Delta_i, X_{t_{i-1}};\theta)] \\ + \sum_{i=1}^{n} \frac{\partial_\theta \sigma^2(X_{t_{i-1}};\theta)}{2\sigma^4(X_{t_{i-1}};\theta)\Delta_i}[(X_{t_i} - F(\Delta_i, X_{t_{i-1}};\theta))^2 - \phi(\Delta_i, X_{t_{i-1}};\theta)]. \tag{6.8}$$

The functions (6.6) and (6.8) are not optimal in the sense of Chapter 4 but they are special cases of so-called *small-Δ-optimal martingale estimating functions*. Small-Δ-optimality is about finding an optimality criterion for martingale estimating functions when the sample size n is fixed and the sampling frequency Δ tends to 0. A concise definition of this concept is given in the next section, where we present some results for ordinary small-Δ-optimal martingale estimating functions. In the subsequent sections, we are going to generalize the existing results by means of the expansions we found in Chapter 5.

6.2 Ordinary small-Delta-optimal martingale estimating functions

In the present section, we rigorously outline some existing results about small-Δ-optimality. These results hold if the underlying diffusion is observed at equidistant time points. This paragraph is based on chapter 6 of Martin Jacobsen's article [37].

Consider the model

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = U,\ t \geq 0, \tag{6.9}$$

where B denotes the Brownian motion of $\mathbb{R}$, and where $\mu : \mathbb{R}\times\Theta \to \mathbb{R}$ and $\sigma : \mathbb{R}\times\Theta \to \mathbb{R}_+$ are sufficiently smooth functions. The parameter θ varies in a subset Θ of $\mathbb{R}$ and the random variable U describes the initial condition. Formally, the diffusion X is defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t)$ with U $\mathcal{F}_0$-measurable. We assume that, for any

$\theta \in \Theta$ and any probability measure ν on $\mathbb{R}$, there is a probability measure $\mathbb{P}_{\nu,\theta}$ on $(\Omega, \mathcal{F})$, with respect to which the σ-algebra $\mathcal{F}_0$ and the Brownian motion B are independent and such that, for the prescribed θ-value, the equation (6.9) has a unique strong solution with ν being the distribution of U.

Recall the result of Theorem 4.2.1.7 which states that on suitable assumptions an estimator $\hat{\theta}_n$ for θ, inferred from a martingale estimating function g, is asymptotically normally distributed. Under the true measure $\mathbb{P}_{\nu,\theta_0}$ we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \longrightarrow N\left(0, \frac{v(\theta_0)}{\xi^2(\theta_0)}\right), \tag{6.10}$$

weakly as the sample size $n \to \infty$. The expressions $v(\theta_0)$ and $\xi(\theta_0)$ are given by $v(\theta_0) = Q^{\Delta}_{\theta_0}(g(\Delta, \cdot; \theta_0)^2)$ and $\xi(\theta_0) = Q^{\Delta}_{\theta_0}(\partial_\theta g(\Delta, \cdot; \theta_0))$, respectively. The probability measure Q^{Δ}_{θ} on $\mathbb{R}^2$ is defined by

$$Q^{\Delta}_{\theta}(x, y) = \nu_\theta(dx) \times p(\Delta, x, y; \theta), \tag{6.11}$$

where $y \mapsto p(\Delta, x, y; \theta)$ denotes the transition density of X_Δ, conditional on $X_0 = x$. In Chapter 4, the aim was to minimize the asymptotic variance $\mathrm{Var}_{\nu,\Delta,\theta}(g, \hat{\theta}) = v(\theta)/\xi^2(\theta)$ for a fixed $\Delta > 0$. By contrast, for the discussion of small-Δ-optimality, we consider $\mathrm{Var}_{\nu,\Delta,\theta}$ for a fixed sample size and for $\Delta \to 0$. We show that in the limit a universal lower bound for the asymptotic variance can be obtained. This implies that, for small values of Δ, an estimator obtained from a small-Δ-optimal estimating function is in practice as good as the maximum likelihood estimator. Thus small-Δ-optimality is a global optimality criterion. Although small-Δ-optimality refers explicitly to the limit $\Delta \to 0$, for any fixed $\Delta > 0$ the estimator is still $\sqrt{n}$-consistent and asymptotically Gaussian as the sample size goes to infinity. There is no guarantee that it is Godambe and Heyde optimal, but for Δ not too large, it should still behave well.

Concretely, expansions to the expressions $v(\theta)$ and $\xi(\theta)$ can be found. A combination of both expansions allows us to expand the asymptotic variance $v(\theta_0)/\xi^2(\theta_0)$ with respect to Δ and this, in turn, enables us to find a universal lower bound. The universal lower bound is simply the first term in the expansion, since higher order terms are asymptotically negligible. The approximate martingale estimating functions in the introduction can then be shown to be optimal in the sense that their variance asymptotically attains the universal lower bound. Presently we will quote the main results of small-Δ-optimality for ordinary martingale estimating functions. But first we have to take some technical assumptions into consideration.

Assumption 6.2.0.8. *The parameter θ belongs to an open subset $\Theta \subset \mathbb{R}$ and for each $\theta \in \Theta$ the diffusion X is ergodic.*

In Section 4.2.1, we briefly discussed conditions on which a diffusion process X, defined by a time-homogenous stochastic differential equation, is ergodic. See Condition 4.2.1.1.

For more details concerning ergodicity and diffusions, see the article of Genon-Catalot et al. [31].

Assumption 6.2.0.9. *The drift and the diffusion coefficient μ and σ of* (6.9) *are supposed to satisfy the following conditions. For all $\theta \in \Theta$, $\mu(x;\theta)$ is continuous in x and, for each $x \in \mathbb{R}$, $\mu(x;\theta)$ is continuously differentiable in θ. The function $\sigma(x,\theta)$ is supposed to be continuously differentiable in (x,θ) and to be uniformly bounded away from 0 in (x,θ).*

We consider a given flow $\mathcal{G} = (g_{\Delta,\theta})_{\Delta\geq 0,\ \theta\in\Theta}$ of well-behaved martingale estimating functions that can be expanded in the following way

$$g_{\Delta,\theta}(x,y) = g_{0,\theta}(x,y) + \Delta\frac{\partial}{\partial s}g_{s,\theta}(x,y)\big|_{s=0} + o(\Delta;\theta,x,y), \tag{6.12}$$

where the notation $o(\Delta;\theta,x,y)$ means that the rest term belongs to $o(\Delta)$ for fixed θ, x, y and has polynomial growth for x, y. We denote with $\mathcal{G}_\theta$ the class of estimating functions g, such that, for all $\Delta \geq 0$ and for all $\theta \in \Theta$, the function $g_{\Delta,\theta}(x,y)$ is continuously differentiable with respect to Δ, continuously differentiable with respect to x and twice continuously differentiable with respect to y. Furthermore, we assume that, for all $g \in \mathcal{G}_\theta$, the following expectations exist

$$\begin{aligned}
&\mathbb{E}_{\nu,\theta}[g^2_{\Delta,\theta}(X_0,X_\Delta)] < \infty,\\
&\mathbb{E}_{\nu,\theta}\left[\left(\frac{\partial}{\partial\Delta}g_{\Delta,\theta}(X_0,X_\Delta) + \mathcal{A}_\theta g_{\Delta,\theta}(X_0,X_\Delta)\right)^2\right] < \infty,\\
&\mathbb{E}_{\nu,\theta}\left[\left(\sigma(X_\Delta;\theta)\frac{\partial}{\partial y}g_{\Delta,\theta}(X_0,y)\Big|_{y=X_\Delta}\right)^2\right] < \infty,
\end{aligned} \tag{6.13}$$

where, for $\theta \in \Theta$, the operator $\mathcal{A}_\theta$ denotes the infinitesimal generator of the diffusion (6.9).

Remark 6.2.0.10. Note that there are different candidates for $\mathcal{G}_\theta$. The most convenient way is to choose $\mathcal{G}_\theta$ as the subclass of martingale estimating functions that consists of functions of the following form

$$g_{\Delta,\theta}(x,y) = \sum_{j=1}^{N} a_j(x,\Delta;\theta)\Big(k_j(y) - \mathbb{E}_{x,\theta}[k_j(X_\Delta)]\Big), \tag{6.14}$$

for different pairs (a_j,k_j), $j = 1,\ldots,N$. Here, the functions a_j and k_j are supposed to satisfy appropriate regularity assumptions with respect to the variables Δ, x, y and with respect to the parameter θ. Finally, note that the linear and the quadratic martingale estimating functions, we mentioned in the introduction of Chapter 4 and in the introduction to this chapter, are contained in the class defined by functions of the type (6.14).

After these preliminary considerations, we are now ready to state the first result. The following proposition presents expansions for the values $\xi(\theta)$ and $v(\theta)$ we introduced above.

Proposition 6.2.0.11. *Let $g_{\Delta,\theta}$, $g^2_{\Delta,\theta}$ and $\frac{\partial}{\partial\theta} g_{\Delta,\theta}$ belong to $\mathcal{G}_\theta$ and assume that g allows for the expansion* (6.12)*. Then*

$$\begin{aligned}
&\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}(X_0, X_\Delta)\right] \\
&= -\Delta \mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta}\mu(X_0,\theta)\frac{\partial}{\partial y} g_{0,\theta}(X_0,y)\big|_{y=X_\Delta} + \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(X_0,\theta)^2\right\}\frac{\partial^2}{\partial y^2} g^2_{0,\theta}(X_0,y)\big|_{y=X_\Delta}\right] \\
&\quad + o(\Delta) \qquad (6.15)
\end{aligned}$$

and

$$\mathbb{E}_{\nu,\theta}\left[g^2_{\Delta,\theta}(X_0, X_\Delta)\right] = \Delta \mathbb{E}_{\nu,\theta}\left[\sigma(X_0,\theta)^2\left\{\frac{\partial}{\partial y} g_{0,\theta}(X_0,y)\big|_{y=X_\Delta}\right\}^2\right] + o(\Delta). \qquad (6.16)$$

Moreover, if $\frac{\partial}{\partial y} g_{0,\theta}(x,y)\big|_{y=x} = 0$, then

$$\mathbb{E}_{\nu,\theta}\left[g^2_{\Delta,\theta}(X_0, X_\Delta)\right] = \frac{1}{2}\Delta^2 \mathbb{E}_{\nu,\theta}\left[\sigma(X_0,\theta)^4\left\{\frac{\partial^2}{\partial y^2} g_{0,\theta}(X_0,y)\big|_{y=X_\Delta}\right\}^2\right] + o(\Delta^2). \qquad (6.17)$$

Proof. See Jacobsen [37], proof of Proposition 5 and the proof of Proposition 6. □

These expansions imply a small-Δ-optimality result, which is stated in the next theorem. To be exact, the theorem presents an expansion of the asymptotic variance $\mathrm{Var}_{\nu,\Delta,\theta}[g,\hat{\theta}]$ with respect to Δ, which allows to infer the asymptotic lower bound.

Theorem 6.2.0.12. *Suppose that $\mathcal{G}$ is a flow of well behaved martingale estimating functions satisfying the expansion* (6.12) *with a non-vanishing $g_{0,\theta}$ such that, for every $\theta \in \Theta$, the functions $\frac{\partial}{\partial\theta} g_{\Delta,\theta}$, $g^2_{\Delta,\theta}$ and $\frac{\partial}{\partial\Delta} g^2_{\Delta,\theta} + \mathcal{A}_\theta g^2_{\Delta,\theta}$ belong to $\mathcal{G}_\theta$.*

(i) If σ^2 does not depend on θ and if the expressions

$$\begin{aligned}
&\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial y} g_{0,\theta}(X_0,y)\big|_{y=X_0}\frac{\partial}{\partial\theta}\mu(X_0;\theta)\right], \\
&\mathbb{E}_{\nu,\theta}\left[\left(\frac{\partial}{\partial\theta}\mu(X_0;\theta)\right)^2\sigma^{-2}(X_0)\right] \qquad (6.18)
\end{aligned}$$

do not vanish, then

$$\mathrm{Var}_{\nu,\Delta,\theta}[g,\hat{\theta}] = \frac{1}{\Delta} v_{-1,\theta}(g,\hat{\theta}) + o\left(\frac{1}{\Delta}\right), \qquad (6.19)$$

where

$$v_{-1,\theta}(g,\hat{\theta}) \geq \left(\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta}\mu(X_0;\theta)\,\sigma^{-2}(X_0)\right]\right)^{-1}. \tag{6.20}$$

Here, equality holds and g is small-Δ-optimal if

$$\frac{\partial}{\partial y}g_{0,\theta}(x,y)\big|_{y=x} \equiv K_\theta \frac{\partial}{\partial\theta}\mu(x;\theta)\sigma^{-2}(x), \tag{6.21}$$

for some constant $K_\theta \neq 0$.

(ii) If σ^2 depends on the parameter θ and if the expressions

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial^2}{\partial y^2}g_{0,\theta}(X_0,y)\big|_{y=X_0}\frac{\partial}{\partial\theta}\sigma(X_0;\theta)\right],$$
$$\mathbb{E}_{\nu,\theta}\left[\left(\frac{\partial}{\partial\theta}\sigma^2(X_0;\theta)\right)^2\sigma^{-4}(X_0;\theta)\right] \tag{6.22}$$

do not vanish, then

$$\mathrm{Var}_{\nu,\Delta,\theta}[g,\hat{\theta}] = \frac{1}{\Delta}v_{-1,\theta}(g,\hat{\theta}) + v_{0,\theta}(g,\hat{\theta}) + o\,(1)\,, \tag{6.23}$$

with $v_{-1,\theta}(g,\hat{\theta}) = 0$ if $\frac{\partial}{\partial y}g_{0,\theta}(x,y)\big|_{y=x} = 0$ for all x and

$$v_{0,\theta}(g,\hat{\theta}) \geq 2\left(\mathbb{E}_{\nu,\theta}\left[\left(\frac{\partial}{\partial\theta}\sigma^2(X_0;\theta)\right)^2\sigma^{-4}(X_0;\theta)\right]\right)^{-1}. \tag{6.24}$$

Here, equality holds and g is small-Δ-optimal if $\frac{\partial}{\partial y}g_{0,\theta}(x,y)\big|_{y=x} \equiv 0$ and

$$\frac{\partial^2}{\partial y^2}g_{0,\theta}(x,y)\big|_{y=x} \equiv K_\theta\,\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\,\sigma(x;\theta)^{-4}, \tag{6.25}$$

for some constant $K_\theta \neq 0$.

In order to prove the theorem, we need an auxiliary result.

Lemma 6.2.0.13. *Let U, Z, S be matrix-valued random variables of dimensions $a \times b$, $a \times b$ and $b \times b$, respectively. The random matrix S is supposed to be symmetric and strictly positive definite with probability 1. Assuming that all entries in the matrices UZ^T, USU^T, $ZS^{-1}Z^T$ are integrable, then the following three properties are satisfied:*

(i) if $\mathbb{E}(USU^T)$ is non-singular, then

$$\mathbb{E}[ZU^T](\mathbb{E}[USU^T])^{-1}\mathbb{E}[UZ^T] \leq \mathbb{E}[ZS^{-1}Z^T], \tag{6.26}$$

(ii) if $\mathbb{E}(USU^T)$ *and* $\mathbb{E}(UZ^T)$ *are non-singular, then*

$$(\mathbb{E}[ZU^T])^{-1}(\mathbb{E}[ZS^{-1}Z^T])(\mathbb{E}[UZ^T])^{-1} \geq (\mathbb{E}[USU^T])^{-1}, \tag{6.27}$$

(iii) in (6.26) *and* (6.27)*, there is equality if for some non-random, non-singular matrix* $K \in \mathbb{R}^{a \times a}$*,* $U = KZS^{-1}$*, equivalently if* $Z = K^{-1}US$*.*

Proof. In the scalar case, *(i)*, *(ii)* and *(iii)* are immediate consequences of Cauchy-Schwarz' inequality. The matrix case is slightly more difficult. An elaborate proof can be found in the appendix of the article [37]. □

Proof of Theorem 6.2.0.12. (i) If σ does not depend on θ, then (see (6.15))

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(X_0, X_\Delta)\right] = -\Delta\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta}\mu(X_0,\theta)\frac{\partial}{\partial y}g_{0,\theta}(X_0,y)\big|_{y=X_\Delta}\right] + o(\Delta). \tag{6.28}$$

Now, apply Lemma 6.2.0.13.

(ii) If $\frac{\partial}{\partial y}g_{0,\theta}(x,y)\big|_{y=x} \equiv 0$, then (see (6.17))

$$\mathbb{E}_{\nu,\theta}\left[g^2_{\Delta,\theta}(X_0, X_\Delta)\right] = \frac{1}{2}\Delta^2\mathbb{E}_{\nu,\theta}\left[\sigma(X_0,\theta)^4\left\{\frac{\partial^2}{\partial y^2}g_{0,\theta}(X_0,y)\big|_{y=X_\Delta}\right\}^2\right] + o(\Delta^2). \tag{6.29}$$

Now, apply Lemma 6.2.0.13. □

Let us briefly sum up the result. The previous theorem states that the variance $\mathrm{Var}_{\nu,\Delta,\theta}(g,\hat{\theta}) = v(\theta)/\xi^2(\theta)$ satisfies the expansion

$$\mathrm{Var}_{\nu,\Delta,\theta}[g,\hat{\theta}] = \frac{1}{\Delta}v_{-1,\theta}(g,\hat{\theta}) + v_{0,\theta}(g,\hat{\theta}) + o\left(1\right), \tag{6.30}$$

as $\Delta \to 0$. Lower bounds for the expressions $v_{-1,\theta}(g,\hat{\theta})$ and $v_{0,\theta}(g,\hat{\theta})$ are given by (6.20) and (6.24). If μ does not depend on θ then $v_{-1,\theta}(g,\hat{\theta})$ vanishes. Thus we obtain asymptotic lower bounds for the variance in case μ depends on θ and in case it does not.

6.3 Generalized small-Delta-optimal martingale estimating functions

The aim of the present section is to find small-Δ-optimality results for generalized martingale estimating functions. In the first paragraph we depict the model we intend to work with. Moreover, we state the assumptions that are necessary for our analysis. In the second paragraph we derive formulae analogous to formula (6.15) and formula (6.16) in Proposition 6.2.0.11 for generalized martingale estimating functions. In the last paragraph we combine these formulae in order to find lower bounds for the variance. This enables us to derive small-Δ-optimality results for several types of generalized martingale estimating functions.

6.3.1 Notation and assumptions

We consider a diffusion process X defined by the stochastic differential equation (6.9). We use the same notations as in Section 6.2 and we assume that Assumption 6.2.0.8 and Assumption 6.2.0.9 are satisfied. Moreover, we impose the following additional assumption.

Assumption 6.3.1.1. *We assume that, if the coefficient σ depends on θ, then*

$$\sigma(x;\theta) = \theta \cdot \sigma(x) \quad \forall \theta \in \Theta,\ \forall x \in \mathbb{R}. \tag{6.31}$$

The previous assumption is necessary, since we are going to consider a transform of X given by

$$Y_t = \theta\sigma(X_0) \int_{y_0}^{X_t} \frac{1}{\theta\sigma(u)} du = \sigma(X_0) \int_{y_0}^{X_t} \frac{1}{\sigma(u)} du, \tag{6.32}$$

for a value $y_0 \in \mathbb{R}$ that has to be specified. Note that Y equals the Lamperti transform of X times the diffusion coefficient $\sigma(X_0;\theta)$ evaluated at X_0. We see that, due to Assumption 6.3.1.1, the Lamperti transform does not depend on the unknown parameter θ. Only because of this assumption, the transform can be calculated. Let F denote the particular primitive $\int_{y_0}^{\cdot} 1/\sigma(u)du$ of $1/\sigma(\cdot)$. In order to simplify our notations, let us assume that the starting point of integration y_0 in (6.32) satisfies

$$y_0 = F^{-1}\left(F(X_0) - \frac{X_0}{\sigma(X_0)}\right). \tag{6.33}$$

Then, the Lamperti transform Y of the process X starting in $X_0 = U$ starts in U as well and it satisfies the following stochastic differential equation

$$\begin{aligned} &dY_t \\ &= \sigma(X_0;\theta)\left(\frac{\mu(Y_t;\theta)}{\sigma(Y_t;\theta)} - \frac{1}{2}\frac{\partial}{\partial y}\sigma(y;\theta)\Big|_{y=Y_t}\right) dt + \sigma(X_0;\theta) dW_t, \quad t \geq 0,\ Y_0 = U. \end{aligned} \tag{6.34}$$

Before we proceed, let us extend our notations. For $\theta \in \Theta$, let $\mathbb{P}_{x,\theta}$ denote the measure for which $\mathbb{P}_{x,\theta}[Y_0 = x] = 1$ and let $\mathbb{E}_{x,\theta}$ denote expectation with respect to this measure. Note that this is consistent with the notation above if $\nu = \delta_x$. Moreover, we define

$$H_t^Y = \sup_{0 \leq s \leq t} Y_s, \quad \text{and} \quad L_t^Y = \inf_{0 \leq s \leq t} Y_s. \tag{6.35}$$

We have to make use of the expansions we derived in the previous chapter. Henceforth, we implicitly assume that the drift coefficient $\mu(x;\theta)$ and the diffusion coefficient $\sigma(x;\theta)$

satisfy, for all $\theta \in \Theta$, the assumptions of Chapter 5 – particularly those of Corollary 5.2.3.6.

We consider a class of flows $\mathcal{G}_\theta = (g_{\Delta,\theta})_{t\geq 0,\theta\in\Theta}$, i.e. a family of functions $g_{\Delta,\theta} : \mathbb{R}^4 \to \mathbb{R}$, parameterized by $(\Delta, \theta) \in \mathbb{R}_+ \times \Theta$. Especially, we restrict ourselves to a subclass of martingale estimating functions that consists of functions of the following form

$$\begin{aligned} &g_{\Delta,\theta}(x,h,l,y) \\ &= \sum_{j=1}^{N} a_j(\Delta, x; \theta) \Big\{ \kappa_j\Big(h - \mathbb{E}_{x,\theta}[H_\Delta],\, l - \mathbb{E}_{x,\theta}[L_\Delta],\, y - \mathbb{E}_{x,\theta}[Y_\Delta]\Big) \\ &\qquad - \mathbb{E}_{x,\theta}\Big[\kappa_j\Big(h - \mathbb{E}_{x,\theta}[H_\Delta],\, l - \mathbb{E}_{x,\theta}[L_\Delta],\, y - \mathbb{E}_{x,\theta}[Y_\Delta]\Big)\Big]\Big\}, \end{aligned} \tag{6.36}$$

for sufficiently smooth functions a_j and κ_j, $j = 1, \ldots, N$. Regularity assumptions on (a_j, κ_j) will be stated in a moment. One might ask why we do not consider functions of the following type

$$g_{\Delta,\theta}(x,h,l,y) = \sum_{j=1}^{N} a_j(\Delta, x; \theta)\Big(\kappa_j(h,l,y) - \mathbb{E}_{x,\theta}[\kappa_j(H_\Delta, L_\Delta, X_\Delta)]\Big). \tag{6.37}$$

The answer is that this would complicate our work. On the other hand, considering (6.36) is not a real constraint compared to working with (6.37).

Before we are going to state the assumptions that our class of estimating functions has to satisfy, we introduce a particular notation. For $g \in \mathcal{G}_\theta$, we set

$$\tilde{g}^{(0)}_{0,\theta}(x,h,l,y) = g_{0,\theta}(x,h,l,y), \tag{6.38}$$

and

$$\tilde{g}^{(1)}_{0,\theta}(x,h,l,y) = \lim_{s\to 0} \frac{1}{\sqrt{s}} \left(g_{s,\theta}(x,h,l,y) - g_{0,\theta}(x,h,l,y)\right). \tag{6.39}$$

And for $n \geq 2$, we set

$$\tilde{g}^{(n)}_{0,\theta}(x,h,l,y) = \lim_{s\to 0} \frac{1}{\sqrt{s^n}} \left(g_{s,\theta}(x,h,l,y) - \sum_{j=0}^{n-1} s^{j/2} \tilde{g}^{(n)}_{0,\theta}(x,h,l,y) \right), \tag{6.40}$$

provided that the limits exists. We will sometimes write $\tilde{\tilde{g}}$ instead of $\tilde{g}^{(2)}$. This definition seems to be strange, therefore we validate it in the following remark.

Remark 6.3.1.2. The functions $\tilde{g}$ are introduced to replace the derivative with respect to the time variable t. For our further analysis we would actually have to apply Itô's formula directly to $g_{\Delta,\theta}(Y_0, H_t^Y, L_t^Y, Y_t)$ in order to obtain

$$\begin{aligned}
\mathbb{E}_{x,\theta}\left[g_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\right] = {} & g_{0,\theta}(x,x,x,x) \\
& + \mathbb{E}_{x,\theta}\left[\int_0^\Delta \frac{\partial}{\partial s} g_{s,\theta}(Y_0, H_s^Y, L_s^Y, Y_s) ds\right] \\
& + \mathbb{E}_{x,\theta}\left[\int_0^\Delta \frac{\partial}{\partial h} g_{s,\theta}(Y_0, h, L_s^Y, Y_s)\Big|_{h=H_s^Y} dH_s^Y\right] \\
& + \mathbb{E}_{x,\theta}\left[\int_0^\Delta \frac{\partial}{\partial l} g_{s,\theta}(Y_0, H_s^Y, l, Y_s)\Big|_{l=L_s^Y} dL_s^Y\right] \\
& + \mathbb{E}_{x,\theta}\left[\int_0^\Delta \frac{\partial}{\partial y} g_{s,\theta}(Y_0, H_s^Y, L_s^Y, y)\Big|_{y=Y_s} dY_s\right] \\
& + \frac{1}{2}\mathbb{E}_{x,\theta}\left[\int_0^\Delta \frac{\partial^2}{\partial y^2} g_{s,\theta}(Y_0, H_s^Y, L_s^Y, y)\Big|_{y=Y_s} d\langle Y\rangle_s\right].
\end{aligned} \tag{6.41}$$

From Chapter 5 we know that, when dealing with running maxima and minima, we have to come to grips with different square root terms in the time variable Δ. Consequently, in the present case it is not trivial to differentiate with respect to Δ. Instead of analyzing (6.41) directly, where g depends on the time variable, we are going to apply Itô's formula to the functions

$$\tilde{g}_{0,\theta}^{(n)}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta), \quad n = 0, 1, 2. \tag{6.42}$$

Then we only have to rearrange the coefficients in the resulting expressions according to the powers in the time variable Δ. The result is a $\sqrt{\Delta}$ power series expansion of the term $\mathbb{E}_{x,\theta}[g_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)]$. Finally, let us note that an equivalent definition of the terms $\tilde{g}^{(1)}$ and $\tilde{g}^{(2)}$ in 6.3.1.3 is given by

$$\begin{aligned}
& g_{t,\theta}(x,h,l,y) \\
& = g_{0,\theta}(x,h,l,y) + \sqrt{t}\lim_{s\to 0}\left(2\sqrt{s}\frac{\partial}{\partial s} g_{s,\theta}(x,h,l,y)\right) \\
& \quad + t\lim_{s\to 0}\left(\frac{\partial}{\partial s} g_{s,\theta}(x,h,l,y) - 2\sqrt{s}\frac{\partial}{\partial s} g_{s,\theta}(x,h,l,y)\right) + O(t^{3/2};\theta,x,h,l,y).
\end{aligned} \tag{6.43}$$

For higher orders, the terms $\tilde{g}^{(k)},\ k \geq 3$, can be derived in a very similar way.

We formulate precise conditions about $\mathcal{G}_\theta$ that we will henceforth assume to be satisfied.

Assumption 6.3.1.3. *The class of flows $\mathcal{G}_\theta$ consists of functions that have the form (6.36). Here, $g \in \mathcal{G}_\theta$ if and only if $\tilde{g}_{0,\theta}$ and $\tilde{g}_{0,\theta}^{(2)}$ exist and if, for all $(x,h,l,y) \in \mathbb{R}^4$, the* *following expansion holds*

$$g_{\Delta,\theta}(x,h,l,y) = g_{0,\theta}(x,h,l,y) + \sqrt{\Delta}\,\tilde{g}_{0,\theta}(x,h,l,y) + \Delta\,\tilde{g}^{(2)}_{0,\theta}(x,h,l,y) \\ + \; O(\Delta^{3/2};\theta,x,h,l,y). \tag{6.44}$$

As above the notation $O(\Delta^{3/2};\theta,x,h,l,y)$ means that the rest term belongs to $O(\Delta^{3/2})$ for fixed (θ,x,h,l,y) and has polynomial growth in the variables x, h, l and y. We assume that, for all $\theta\in\Theta$ and $\Delta\geq 0$, the function $g_{\Delta,\theta}(x,h,l,y)$ is continuous in x and 3 times continuously differentiable with respect to each variable h, l and y. Correspondingly, we assume that, for all $\theta\in\Theta$, also the functions $g^{(0)}_{0,\theta}(x,h,l,y)$, $\tilde{g}^{(1)}_{0,\theta}(x,h,l,y)$ and $\tilde{g}^{(2)}_{0,\theta}(x,h,l,y)$ are continuous in x and 3 times continuously differentiable in (h,l,y).

It is necessary to comment on these assumptions. Assumption 6.3.1.3 is reasonable, since the class of martingale estimating functions we consider, consists of functions g having the specific form (6.36), with real valued functions κ_j, $j=1,\ldots,N$, that basically behave like polynomials in

$$H_\Delta - \mathbb{E}_{x,\theta}[H_\Delta], \quad L_\Delta - \mathbb{E}_{x,\theta}[L_\Delta] \quad \text{and} \quad Y_\Delta - \mathbb{E}_{x,\theta}[Y_\Delta]. \tag{6.45}$$

Hence, the results of Chapter 5 enable us to find an expansion of g with respect to $\sqrt{\Delta}$ according to (6.44), provided the weight terms a_j, $j=1,\ldots,N$, in formula (6.36) behave decently.

Lastly, let us introduce the following notations. For a sufficiently smooth function $g:\mathbb{R}^4\to\mathbb{R}$ that does not depend on the time variable Δ and that satisfies a polynomial growth condition, we have the expansion

$$\mathbb{E}_{x,\theta}[g(Y_0,H^Y_\Delta,L^Y_\Delta,Y_\Delta)] = g(x,x,x,x) + \sqrt{\Delta}\mathcal{A}^{(\frac{1}{2})}_\theta g(x,x,x,x) + \Delta\mathcal{A}^{(1)}_\theta g(x,x,x,x) \\ + \; O(\Delta^{3/2}), \tag{6.46}$$

where the operators $\mathcal{A}^{(1/2)}_\theta$ and $\mathcal{A}^{(1)}_\theta$ are given by

$$\mathcal{A}^{(\frac{1}{2})}_\theta g(x,x,x,x) = \sigma(x;\theta)\frac{2}{\sqrt{2\pi}}g_{0,1,0,0}(x,x,x,x) - \sigma(x;\theta)\frac{2}{\sqrt{2\pi}}g_{0,0,1,0}(x,x,x,x) \tag{6.47}$$

and

$$\begin{aligned}
&\mathcal{A}^{(1)}_\theta g(x,x,x,x)\\
&= g_{0,1,0,0}(x,x,x,x)\frac{1}{2}\Big(\mu(x;\theta) - \frac{1}{2}\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\Big) + g_{0,2,0,0}(x,x,x,x)\frac{1}{2}\sigma^2(x;\theta)\\
&\;+ g_{0,0,1,0}(x,x,x,x)\frac{1}{2}\Big(\mu(x;\theta) - \frac{1}{2}\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\Big) + g_{0,0,2,0}(x,x,x,x)\frac{1}{2}\sigma^2(x;\theta)\\
&\;+ (1-2\log 2)g_{0,1,1,0}(x,x,x,x)\sigma^2(x;\theta)
\end{aligned}$$

$$+\frac{1}{2}g_{0,1,0,1}(x,x,x,x)\sigma^2(x;\theta)+\frac{1}{2}g_{0,0,1,1}(x,x,x,x)\sigma^2(x;\theta)$$
$$+g_{0,0,0,1}(x,x,x,x)\left(\mu(x;\theta)-\frac{1}{2}\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right)+\frac{1}{2}g_{0,0,0,2}(x,x,x,x)\sigma^2(x;\theta), \tag{6.48}$$

respectively. Here, for a multi-index $\alpha \in \mathbb{N}_0^4$, g_α denotes the derivative described in (5.81).

6.3.2 Auxiliary results - expansions

The aim of the present paragraph is to find formulae analogous to (6.15) and (6.16) in Proposition 6.2.0.11 for generalized martingale estimating functions. Concretely, we derive a second order expansions for both

$$\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta)\right] \tag{6.49}$$

and

$$\mathbb{E}_{x,\theta}\left[g_{\Delta,\theta}^2(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta)\right]. \tag{6.50}$$

with respect to $\sqrt{\Delta}$. The second order expansion of (6.50) holds in general. But for a particular type of estimating functions $g_{\Delta,\theta}$ we are even able to infer a fourth order expansion, with respect to $\sqrt{\Delta}$, for this expression. Eventually, note that the respective expansions for ordinary martingale estimating functions can be recovered from the results stated below. Therefore, our expansions can also be considered as a generalization of the expansions in Proposition 6.2.0.11.

We start with a technical result that will turn out to be crucial in the sequel.

Proposition 6.3.2.1. *Let Y denote the process defined by the stochastic differential equation* (6.34) *and let $\mathcal{G}_\theta$ be a class of flows that satisfies Assumption 6.3.1.3. Then, for $g \in \mathcal{G}_\theta$, we have*

$$\begin{aligned}0 &= \mathbb{E}_{x,\theta}[g_{\Delta,\theta}(Y_0,H_t^Y,L_t^Y,Y_t)]\\ &= g_{0,\theta}(x,x,x,x)+\sqrt{\Delta}\mathcal{A}_\theta^{(\frac{1}{2})}g_{0,\theta}(x,x,x,x)+\sqrt{\Delta}\tilde{g}_{0,\theta}(x,x,x,x)\\ &\quad+\Delta\mathcal{A}_\theta^{(\frac{1}{2})}\tilde{g}_{0,\theta}(x,x,x,x)\\ &\quad+\Delta\mathcal{A}_\theta^{(1)}g_{0,\theta}(x,x,x,x)+\Delta\tilde{\tilde{g}}_{0,\theta}(x,x,x,x)+O(\Delta^{3/2}),\end{aligned} \tag{6.51}$$

where the operators $\mathcal{A}_\theta^{(\frac{1}{2})}$ and $\mathcal{A}_\theta^{(1)}$ are given by (6.47) *and* (6.48)*, respectively.*

Proof. By Assumption 6.3.1.3 the function g has an expansion with respect to $\sqrt{\Delta}$ of the following form

$$\begin{aligned} &g_{\Delta,\theta}(x,h,l,y) \\ &= g_{0,\theta}(x,h,l,y) + \sqrt{\Delta}\tilde{g}_{0,\theta}(x,h,l,y) + \Delta\tilde{\tilde{g}}_{0,\theta}(x,h,l,y) + O(\Delta^{3/2};\,\theta,x,h,l,y). \end{aligned} \tag{6.52}$$

In view of the fact that the rest term $O(\Delta^{3/2};\,\theta,x,h,l,y)$ behaves like a polynomial in h, l and y, the result can be inferred by an expansion of each of the terms

$$\mathbb{E}_x\Big[\tilde{g}^{(k)}_{0,\theta}(Y_0,H^Y_t,L^Y_t,Y_t)\Big], \quad k=0,1,2. \tag{6.53}$$

A comparison with Corollary 5.2.3.6 yields the result. □

Let us analyze formula (6.51) a little further. By letting $\Delta \to 0$, we find

$$g_{0,\theta}(x,x,x,x) = 0. \tag{6.54}$$

And moreover, we easily infer that

$$\mathcal{A}^{(\frac{1}{2})}_\theta g_{0,\theta}(x,x,x,x) + \tilde{g}_{0,\theta}(x,x,x,x) = 0, \tag{6.55}$$

$$\mathcal{A}^{(\frac{1}{2})}_\theta \tilde{g}_{0,\theta}(x,x,x,x) + \mathcal{A}^{(1)}_\theta g_{0,\theta}(x,x,x,x) + \tilde{\tilde{g}}_{0,\theta}(x,x,x,x) = 0. \tag{6.56}$$

Note that equation (6.56) is equivalent to

$$-\left(\mathcal{A}^{(\frac{1}{2})}_\theta\right)^2 g_{0,\theta}(x,x,x,x) + \mathcal{A}^{(1)}_\theta g_{0,\theta}(x,x,x,x) + \tilde{\tilde{g}}_{0,\theta}(x,x,x,x) = 0, \tag{6.57}$$

where the operator $\left(\mathcal{A}^{(\frac{1}{2})}_\theta\right)^2$ is given by

$$\begin{aligned} \left(\mathcal{A}^{(\frac{1}{2})}_\theta\right)^2 g_{0,\theta}(x,x,x,x) &= \sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \\ &+\ \sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial l^2} g_{0,\theta}(x,x,l,x)\Big|_{l=x} \\ &-\ 2\sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial h \partial l} g_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x}. \end{aligned} \tag{6.58}$$

This follows directly by means of equation (6.55).
On account of its importance we summarize this result in the following corollary.

Corollary 6.3.2.2. *Let the assumptions of Proposition 6.3.2.1 be satisfied. Then*

$$\begin{aligned} 0 &= \mathbb{E}_{x,\theta}[g_{\Delta,\theta}(Y_0,H^Y_t,L^Y_t,Y_t)] \\ &= g_{0,\theta}(x,x,x,x) + \sqrt{\Delta}\mathcal{A}^{(\frac{1}{2})}_\theta g_{0,\theta}(x,x,x,x) + \sqrt{\Delta}\tilde{g}_{0,\theta}(x,x,x,x) \end{aligned}$$

$$
\begin{aligned}
&- \Delta \left(\mathcal{A}_\theta^{(\frac{1}{2})} \right)^2 \tilde{g}_{0,\theta}(x,x,x,x) \\
&+ \Delta \mathcal{A}_\theta^{(1)} g_{0,\theta}(x,x,x,x) + \Delta \tilde{\tilde{g}}_{0,\theta}(x,x,x,x) + O(\Delta^{3/2}).
\end{aligned} \tag{6.59}
$$

Proof. The result follows directly from formulae (6.55) and (6.57). □

We are now in a position to derive our first expansion. The next proposition states the analogue of formula (6.15) in Proposition 6.2.0.11 for generalized martingale estimating functions.

Proposition 6.3.2.3. *Let the assumptions of Proposition 6.3.2.1 be satisfied. Then*

$$
\begin{aligned}
\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\right] = &- \sqrt{\Delta}\, \dot{\mathcal{A}}_\theta^{(\frac{1}{2})} g_{0,\theta}(x,x,x,x) + \dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})} \right)^2} g_{0,\theta}(x,x,x,x) \\
&- \Delta \dot{\mathcal{A}}_\theta^{(1)} g_{0,\theta}(x,x,x,x) + O(\Delta^{3/2}),
\end{aligned} \tag{6.60}
$$

where the operators $\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}$ *and* $\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})} \right)^2}$ *are given by*

$$
\begin{aligned}
\dot{\mathcal{A}}_\theta^{(\frac{1}{2})} g_{0,\theta}(x,x,x,x) = &\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\frac{\partial}{\partial h} g_{0,\theta}(x,x,h,x)\Big|_{h=x} \\
&- \frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\frac{\partial}{\partial l} g_{0,\theta}(x,x,l,x)\Big|_{l=x},
\end{aligned} \tag{6.61}
$$

and

$$
\begin{aligned}
\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})} \right)^2} g_{0,\theta}(x,x,x,x) = &\frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \\
&+ \frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial l^2} g_{0,\theta}(x,x,l,x)\Big|_{l=x} \\
&- \frac{\partial}{\partial\theta}\sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial h \partial l} g_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x},
\end{aligned} \tag{6.62}
$$

respectively. Finally, $\dot{\mathcal{A}}^{(1)}$ *is given by*

$$
\begin{aligned}
\dot{\mathcal{A}}_\theta^{(1)} g_{0,\theta}(x,x,x,x) = &\frac{1}{2}\left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right)\frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \\
&+ \frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \\
&+ \frac{1}{2}\left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right)\frac{\partial}{\partial l} g_{0,\theta}(x,x,l,x)\Big|_{l=x} \\
&+ \frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\frac{\partial^2}{\partial l^2} g_{0,\theta}(x,x,l,x)\Big|_{l=x} \\
&+ (1 - 2\log 2)\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\frac{\partial^2}{\partial h \partial l} g_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\frac{\partial^2}{\partial h\partial y}g_{0,\theta}(x,h,x,y)\Big|_{h=x,y=x}\\
&+\frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\frac{\partial^2}{\partial l\partial y}g_{0,\theta}(x,x,l,y)\Big|_{l=x,y=x}\\
&+\left(\frac{\partial}{\partial\theta}\mu(x;\theta)-\frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right)\frac{\partial}{\partial y}g_{0,\theta}(x,x,x,y)\Big|_{y=x}\\
&+\frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2\frac{\partial^2}{\partial y^2}g_{0,\theta}(x,x,x,y)\Big|_{y=x}. \qquad (6.63)
\end{aligned}
$$

Remark 6.3.2.4. If the diffusion coefficient σ of the diffusion (6.9) does not depend on the parameter θ, the operator $\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}$ vanishes, and we have the following result

$$
\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta)\right]=\Delta\dot{\mathcal{A}}_\theta^{(1)}g_{0,\theta}(x,x,x,x)+O(\Delta^{3/2}). \qquad (6.64)
$$

The operator $\dot{\mathcal{A}}^{(1)}$ also takes a simple form in this case, namely

$$
\begin{aligned}
\dot{\mathcal{A}}_\theta^{(1)}g_{0,\theta}(x,x,x,x)&=\frac{1}{2}\frac{\partial}{\partial\theta}\mu(x;\theta)\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\\
&+\frac{1}{2}\frac{\partial}{\partial\theta}\mu(x;\theta)\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}\\
&+\frac{\partial}{\partial\theta}\mu(x;\theta)\frac{\partial}{\partial y}g_{0,\theta}(x,x,x,y)\Big|_{y=x}. \qquad (6.65)
\end{aligned}
$$

A similar result can be stated if μ does not depend on θ. In any case, we have the first order expansion

$$
\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta)\right]=\sqrt{\Delta}\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}g_{0,\theta}(x,x,x,x)+O(\Delta). \qquad (6.66)
$$

Proof (of Proposition 6.3.2.3). We have

$$
\begin{aligned}
&\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta)\right]\\
&=\Bigg\{\frac{\partial}{\partial\theta}g_{0,\theta}+\sqrt{\Delta}\mathcal{A}_\theta^{(\frac{1}{2})}\frac{\partial}{\partial\theta}g_{0,\theta}+\sqrt{\Delta}\frac{\partial}{\partial\theta}\tilde{g}_{0,\theta}\\
&\qquad+\Delta\mathcal{A}^{(\frac{1}{2})}\frac{\partial}{\partial\theta}\tilde{g}_{0,\theta}+\Delta\mathcal{A}_\theta^{(1)}\frac{\partial}{\partial\theta}g_{0,\theta}+\Delta\frac{\partial}{\partial\theta}\tilde{\tilde{g}}_{0,\theta}\Bigg\}(x,x,x,x)+O(\Delta^{3/2}). \qquad (6.67)
\end{aligned}
$$

Now, $\frac{\partial}{\partial\theta}g_{0,\theta}(x,x,x,x)\equiv 0$ and

$$
\begin{aligned}
0&=\frac{\partial}{\partial\theta}\Big(\mathcal{A}_\theta^{(\frac{1}{2})}g_{0,\theta}(x,x,x,x)+\tilde{g}_{0,\theta}(x,x,x,x)\Big)\\
&=\Big(\mathcal{A}_\theta^{(\frac{1}{2})}\frac{\partial}{\partial\theta}g_{0,\theta}(x,x,x,x)+\frac{\partial}{\partial\theta}\tilde{g}_{0,\theta}(x,x,x,x)\Big)+\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}g_{0,\theta}(x,x,x,x), \qquad (6.68)
\end{aligned}
$$

where $\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}$ is the operator (6.61). Analogously, we find that

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\Big(- \left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 g_{0,\theta}(x,x,x,x) + \mathcal{A}_\theta^{(1)} g_{0,\theta}(x,x,x,x) + \tilde{\tilde{g}}_{0,\theta}(x,x,x,x)\Big) \\
&= \Big(- \left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 \frac{\partial}{\partial\theta} g_{0,\theta}(x,x,x,x) + \mathcal{A}_\theta^{(1)} \frac{\partial}{\partial\theta} g_{0,\theta}(x,x,x,x) + \frac{\partial}{\partial\theta}\tilde{\tilde{g}}_{0,\theta}(x,x,x,x)\Big) \\
&\quad - \dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2} g_{0,\theta}(x,x,x,x) + \dot{\mathcal{A}}_\theta^{(1)} g_{0,\theta}(x,x,x,x),
\end{aligned} \tag{6.69}$$

where $\dot{\mathcal{A}}^{(1)}$ is given by (6.63). Since

$$-\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2} g_{0,\theta}(x,x,x,x) = \dot{\mathcal{A}}_\theta^{(\frac{1}{2})} \tilde{g}_{0,\theta}(x,x,x,x), \tag{6.70}$$

inserting (6.68) and (6.69) into equation (6.67) yields the result. □

Now that we have sufficiently analyzed the expansion of (6.49), let us move on to the analysis of (6.50). The next result states the analogue of formula (6.16) in Proposition 6.2.0.11.

Proposition 6.3.2.5. *Let the assumptions of Proposition 6.3.2.1 be satisfied. Then*

$$\mathbb{E}_{x,\theta}\Big[g_{\Delta,\theta}^2(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\Big] = \Delta \mathcal{A}^S g_{0,\theta}(x,x,x,x) + O(\Delta^{3/2}), \tag{6.71}$$

where the operator $\mathcal{A}^S$ is defined via

$$\begin{aligned}
&\mathcal{A}^S g_{0,\theta}(x,x,x,x) \\
&= \left(1 - \frac{2}{\pi}\right)\left(\frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x}\right)^2 \sigma(x;\theta)^2 \\
&\quad + \left(1 - \frac{2}{\pi}\right)\left(\frac{\partial}{\partial l} g_{0,\theta}(x,x,l,x)\Big|_{l=x}\right)^2 \sigma(x;\theta)^2 \\
&\quad + \left(\frac{4}{\pi} + 2\left(1 - 2\log 2\right)\right)\left(\frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \frac{\partial}{\partial l} g_{0,\theta}(x,x,l,x)\Big|_{l=x}\right) \sigma(x;\theta)^2 \\
&\quad + \left(\frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \frac{\partial}{\partial y} g_{0,\theta}(x,x,x,y)\Big|_{y=x}\right) \sigma(x;\theta)^2 \\
&\quad + \left(\frac{\partial}{\partial l} g_{0,\theta}(x,h,x,x)\Big|_{l=x} \frac{\partial}{\partial y} g_{0,\theta}(x,x,x,y)\Big|_{y=x}\right) \sigma(x;\theta)^2 \\
&\quad + \left(\frac{\partial}{\partial y} g_{0,\theta}(x,x,x,y)\Big|_{y=x}\right)^2 \sigma(x;\theta)^2.
\end{aligned} \tag{6.72}$$

Proof. We moved the proof to Appendix 10.2. □

Note that the case

$$\frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x} = \frac{\partial}{\partial l} g_{0,\theta}(x,x,l,x)\Big|_{l=x} = \frac{\partial}{\partial y} g_{0,\theta}(x,x,x,y)\Big|_{h=y} = 0 \qquad (6.73)$$

is not covered by the results we have stated so far. A further expansion of the martingale estimating function with respect to $\sqrt{\Delta}$ is required. The special form of our martingale estimating functions allows us to determine a fourth order expansion if condition (6.73) holds. For the sake of simplicity it will only be derived for the case, where the martingale estimating function g depends on the variables h, y and is independent of the minimum variable l.

Proposition 6.3.2.6. *Let Y denote the process defined by the stochastic differential equation* (6.34)*. We assume that $\mathcal{G}_\theta$ is a class of flows that satisfies Assumption 6.3.1.3. Then for any $g \in \mathcal{G}_\theta$, that satisfies the additional assumption* (6.73)*, the following expansion holds:*

$$\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, Y_\Delta)\Big] = \Delta^2 \mathcal{A}^S_2 g_{0,\theta}(x,x,x) + O(\Delta^{5/2}), \qquad (6.74)$$

where the operator $\mathcal{A}^S_2$ is defined via

$$\begin{aligned}
&\mathcal{A}^S_2 g_{0,\theta}(x,x,x)\\
&= \sigma(x;\theta)^4 \Bigg\{ \left(\frac{1}{2} - \frac{4}{\pi^2}\right) \left(\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x)\Big|_{h=x} \right)^2 \\
&+ \frac{1}{2} \left(\frac{\partial^2}{\partial y^2} g_{0,\theta}(x,x,y)\Big|_{y=x} \right)^2 \\
&+ \frac{\partial}{\partial h}\frac{\partial}{\partial y} g_{0,\theta}(x,h,y)\Big|_{h=x,y=x} \frac{\partial^2}{\partial y^2} g_{0,\theta}(x,x,y)\Big|_{y=x} \\
&+ \left(\frac{7}{4} - \frac{4}{\pi}\right) \frac{\partial}{\partial h}\frac{\partial}{\partial y} g_{0,\theta}(x,h,y)\Big|_{h=x,y=x} \frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x)\Big|_{y=x} \\
&+ \left(\frac{7}{4} - \frac{10}{3\pi}\right) \left(\frac{\partial}{\partial h}\frac{\partial}{\partial y} g_{0,\theta}(x,h,y)\Big|_{h=x,y=x} \right)^2 \\
&+ \left(\frac{1}{2} - \frac{2}{3\pi}\right) \frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x)\Big|_{h=x} \frac{\partial^2}{\partial y^2} g_{0,\theta}(x,x,y)\Big|_{y=x} \Bigg\}. \qquad (6.75)
\end{aligned}$$

Proof. We moved the proof to Appendix 10.2. □

Before considering concrete examples, let us state some remarks.

Remark 6.3.2.7. Especially, the last result allows us to treat martingale estimating functions given by

$$g(\Delta, x, h, y; \theta) = \sum_{j=1}^{N} a_j(\Delta, x; \theta) k_j(\Delta, x, h, y; \theta) \tag{6.76}$$

with functions k_j that have the form

$$\begin{aligned} k_j(\Delta, x, h, y; \theta) = {} & [h - F^H(\Delta, x; \theta)]^m [y - F^Y(\Delta, x; \theta)]^n \\ & - \mathbb{E}_{x,\theta}\Big[[H_\Delta - F^H(\Delta, x; \theta)]^m \, [Y_\Delta - F^Y(\Delta, x; \theta)]^n \Big], \end{aligned} \tag{6.77}$$

for $m, n \in \mathbb{N}$ with $m + n \geq 2$. Clearly in this situation, the estimate (6.74) holds. Particularly, in the next section, we will be interested in the estimating function

$$g_{qua}(\Delta, x, h, y; \theta) = \sum_{j=1}^{3} a_j(\Delta, x; \theta) k_j(\Delta, x, h, y; \theta) \tag{6.78}$$

with

$$\begin{aligned} k_1(\Delta, x, h, y; \theta) &= [h - F^H(\Delta, x; \theta)]^2 - \phi_{H,H}(\Delta, x; \theta), \\ k_2(\Delta, x, h, y; \theta) &= [y - F^Y(\Delta, x; \theta)]^2 - \phi_{Y,Y}(\Delta, x; \theta), \\ k_3(\Delta, x, h, y; \theta) &= [h - F^H(\Delta, x; \theta)][y - F^Y(\Delta, x; \theta)] - \phi_{H,Y}(\Delta, x; \theta), \end{aligned} \tag{6.79}$$

where F denotes

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[U], \quad \text{for } U \in \{Y_\Delta, H_\Delta^Y\}, \tag{6.80}$$

and ϕ denotes

$$\phi_{U,V}(\Delta, x; \theta) = \mathrm{Cov}_{x,\theta}[U, V], \quad \text{for } U, V \in \{Y_\Delta, H_\Delta^Y\}. \tag{6.81}$$

Remark 6.3.2.8. For a martingale estimating function

$$g(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{N} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta) \tag{6.82}$$

that depends on all variables h, l and y, where the functions k_j are given by

$$\begin{aligned} k_j(\Delta, x, h, l, y; \theta) = {} & \Big\{ \kappa_j\Big(h - \mathbb{E}_{x,\theta}[H_\Delta], l - \mathbb{E}_{x,\theta}[L_\Delta], y - \mathbb{E}_{x,\theta}[Y_\Delta]\Big) \\ & - \mathbb{E}_{x,\theta}\Big[\kappa_j\Big(h - \mathbb{E}_{x,\theta}[H_\Delta], l - \mathbb{E}_{x,\theta}[L_\Delta], y - \mathbb{E}_{x,\theta}[Y_\Delta]\Big)\Big]\Big\}, \end{aligned} \tag{6.83}$$

a result similar to the one of Proposition 6.3.2.6 can be found, provided condition (6.73) is satisfied. In this case the differential operator $\mathcal{A}_2^S$ depends on fourth order moments of $(H^Y_\Delta, L^Y_\Delta, Y_\Delta)$ or rather on the corresponding moments of Brownian motion. But only the value $\mathbb{E}_x[H^B_1 \cdot L^B_1] = 1 - 2\log 2$ is know explicitly. Higher order moments of the vector $(H^B_1, L^B_1) = \big(\sup_{0\le s\le 1} B_s,\, \inf_{0\le s\le 1} B_s\big)$ can only be calculated numerically. We gave an approximation to some moments in Table 5.1 in Remark 5.2.1.5. Although the calculations are straightforward, we are not going to display the overall result here. We will not consider this particular example in the sequel.

Remark 6.3.2.9. The proof of Proposition 6.3.2.6 relies heavily on the fact that we deal with an estimating function that has quadratic terms only. Note that the proof does not work any more for martingale estimating functions that have both linear and quadratic terms, since in this case condition (6.73) is violated. To capture the particularities of such a model, an expansion of $\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, L^Y_\Delta, Y_\Delta)\Big]$ with respect to $\sqrt{\Delta}$ including powers of 3 is required. So far, we have not been able to calculate an approximation of this order. In Chapter 7 we will present the tools for calculating an overall expansion of $\mathbb{E}_{x,\theta}\Big[\gamma(X_0, H_\Delta, X_\Delta)\Big]$ with respect to $\sqrt{\Delta}$ for a certain class of diffusions X and for sufficiently smooth functions $\gamma : \mathbb{R}^3 \to \mathbb{R}$. Even though we are not going to do this in the present thesis, the overall expansion can obviously be used to derive more sophisticated results about small-Δ-optimality for generalized martingale estimating functions based on the observation (H_Δ, X_Δ).

6.3.3 Generalized small-Delta-optimal martingal estimating functions

In this section we present small-Δ-optimality results for several classes of generalized martingale estimating functions. The theoretical tools for deriving a lower bound for the variance of generalized martingale estimating functions were proved in the previous paragraph. Throughout this section, we consider flows $\mathcal{G}_\theta = (g_{\Delta,\theta})_{\Delta\ge 0,\ \theta\in\Theta}$ of well-behaved estimating functions, that satisfy Assumption 6.3.1.3.

Let us consider a model X defined by (6.1). We assume that X starts in $X_0 = U$ and we suppose that ν denotes the law of U. Let Y denote the process defined by (6.34). For a flow $(g_{\Delta,\theta})_{\Delta\ge 0,\theta\in\Theta}$, set $v(\theta) = Q^\Delta_\theta(g_{\Delta,\theta}(\cdot)^2)$ and $\xi(\theta) = Q^\Delta_\theta(\partial_\theta g_{\Delta,\theta}(\cdot))$, where the probability measure Q^Δ_θ on $\mathbb{R}^4$ is defined by

$$Q^\Delta_\theta(x,h,l,y) = \nu_\theta(dx) \times f^Y(\Delta, x, h, l, y; \theta), \tag{6.84}$$

and $(h,l,y) \mapsto f^Y(\Delta, x, h, l, y; \theta)$ denotes the joint density of the triplet

$$(H^Y_\Delta, L^Y_\Delta, Y_\Delta) = \left(\sup_{0\le t\le\Delta} Y_t, \inf_{0\le t\le\Delta} Y_t, Y_\Delta\right), \tag{6.85}$$

conditional on $Y_0 = x$. Finally, we define

$$\mathrm{Var}_{\Delta,\nu,\theta}[g,\hat{\theta}] = \frac{v(\theta)}{\xi(\theta)^2}. \tag{6.86}$$

Compare also the statement of Theorem 4.2.1.7.

A linear estimator for the drift

Recall the results of the Propositions 6.3.2.3 and 6.3.2.5, where we found expansions for

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\right] \tag{6.87}$$

and

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\Big]. \tag{6.88}$$

For $x \in \mathbb{R}$, let $Z_{lin}(x)$ be the vector

$$Z_{lin}(x) = \left(\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}, \frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}, \frac{\partial}{\partial y}g_{0,\theta}(x,x,x,y)\Big|_{y=x}\right), \tag{6.89}$$

and let $U_{lin}^{(1/2)}$ be the vector

$$\begin{aligned} U_{lin}^{(1/2)}(x) &= \left(U_{lin,1}^{(1/2)}(x), U_{lin,2}^{(1/2)}(x), U_{lin,3}^{(1/2)}(x)\right) \\ &= \left(\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}, -\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}, 0\right). \end{aligned} \tag{6.90}$$

Note that $U_{lin}^{(1/2)}$ is zero if the diffusion coefficient of the underlying process X given by (6.9) does not depend on θ. Furthermore, we define $U_{lin}^{(1)}$ to be the vector

$$U_{lin}^{(1)}(x) = \left(U_{lin,1}^{(1)}(x), U_{lin,2}^{(1)}(x), U_{lin,3}^{(1)}(x)\right), \ x \in \mathbb{R}, \tag{6.91}$$

having the entries

$$U_{lin,1}^{(1)}(x) = U_{lin,2}^{(1)}(x) = \frac{1}{2}\left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right) \tag{6.92}$$

and

$$U_{lin,3}^{(1)}(x) = \left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right). \tag{6.93}$$

Moreover, let S_{lin}^{-1} be the matrix

$$S_{lin}^{-1}(x) = \theta^2\sigma^2(x)\begin{pmatrix} 1-\frac{2}{\pi} & \frac{2}{\pi}+(1-2\log 2) & \frac{1}{2} \\ \frac{2}{\pi}+(1-2\log 2) & 1-\frac{2}{\pi} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}, \ x \in \mathbb{R}, \tag{6.94}$$

and S_{lin} its inverse

$$S_{lin}(x) = \frac{1}{\theta^2\sigma^2(x)} \begin{pmatrix} \frac{8-3\pi}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{8+\pi(3-8\log 2)}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{1}{-3+\log 16} \\ \frac{8+\pi(3-8\log 2)}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{8-3\pi}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{1}{-3+\log 16} \\ \frac{1}{-3+\log 16} & \frac{1}{-3+\log 16} & \frac{-4+\log 16}{-3+\log 16} \end{pmatrix}, \ x \in \mathbb{R}. \tag{6.95}$$

The matrix S_{lin}^{-1} is positive definite for all $x \in \mathbb{R}$. This follows from the fact that S_{lin}^{-1} is a covariance matrix. The previous formulae enable us to rewrite the expansions for the derivative of the martingale estimating function g with respect to θ and the expansion for g^2. First, we have

$$\begin{aligned} &\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\right] \\ &= \sqrt{\Delta}\mathbb{E}_{\nu,\theta}[Z_{lin}(\cdot)U_{lin}^{(1/2)}(\cdot)^T] + \Delta\mathbb{E}_{\nu,\theta}[Z_{lin}(\cdot)U_{lin}^{(1)}(\cdot)^T] + o(\Delta) \\ &= \mathbb{E}_{\nu,\theta}[Z_{lin}U_{lin}^T(\Delta,\cdot)] + o(\Delta), \end{aligned} \tag{6.96}$$

where we set

$$U_{lin}(\Delta, x) = \sqrt{\Delta}U_{lin}^{(1/2)}(x) + \Delta U_{lin}^{(1)}(x), \ x \in \mathbb{R}. \tag{6.97}$$

And furthermore, we have

$$\mathbb{E}_{\nu,\theta}\Big[g_{\Delta,\theta}^2(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta)\Big] = \Delta\mathbb{E}_{\nu,\theta}[Z_{lin}(\cdot)S_{lin}^{-1}(\cdot)Z_{lin}(\cdot)^T] + o(\Delta). \tag{6.98}$$

The statement of Theorem 4.2.1.7 in combination with Lemma 6.2.0.13 gives a lower bound for the variance which will be described in the next theorem.

Theorem 6.3.3.1. *Suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 6.3.1.3. If both coefficients μ and σ of the underlying process X depend on θ, then for all $g \in \mathcal{G}_\theta$*

$$\mathrm{Var}_{\Delta,\nu,\theta}[g,\hat{\theta}] = \mathcal{V}_{-1,\theta}(\Delta, g, \hat{\theta}) + o\left(\frac{1}{\Delta}\right), \tag{6.99}$$

where

$$\begin{aligned} \mathcal{V}_{-1,\theta}(\Delta, g, \hat{\theta}) &\geq \Delta\Big(\mathbb{E}_{\nu,\theta}\left[U_{lin}(\Delta,\cdot)\, S_{lin}(\cdot)\, U_{lin}(\Delta,\cdot)^T\right]\Big)^{-1} \\ &= \Delta\left(\mathbb{E}_{\nu,\theta}\left[\Delta\left(\Delta U_{lin}^{(1)}(\cdot)^2 + \frac{\pi U_{lin}^{(1/2)}(\cdot)^2}{\pi\log 2 - 2}\right)\right]\right)^{-1}, \end{aligned} \tag{6.100}$$

provided that the participating expectations do not vanish. In (6.100) *equality holds, and consequently g is small-Δ-optimal if there is a scalar $K \in \mathbb{R}$, possibly depending on θ,* *such that*

$$Z_{lin} = K\, U_{lin}(\Delta, \cdot) S_{lin}. \tag{6.101}$$

Proof. Recall the definition of $\mathrm{Var}_{\Delta,\nu,\theta}[g, \hat{\theta}]$ in (6.86). Division of (6.98) by (6.96) squared, in combination with the inequalities of Lemma 6.2.0.13, yields the result. □

We discuss the latter result a little further. Inequality (6.100) is not very handy, since the lower bound of $\mathcal{V}_{-1,\theta}(\Delta, g, \hat{\theta})$ depends on Δ in an unpleasant way. But let us consider a special case. If the diffusion coefficient σ does not depend on the parameter θ, the vector $U_{lin}^{(1/2)}$ is identically equal to zero and $U_{lin}^{(1)}(x)$ becomes

$$\begin{aligned} U_{lin}^{(1)}(x) &= (U_{lin,1}^{(1)}(x), U_{lin,2}^{(1)}(x), U_{lin,3}^{(1)}(x)) \\ &= \left(\frac{1}{2}\frac{\partial}{\partial\theta}\mu(x;\theta), \frac{1}{2}\frac{\partial}{\partial\theta}\mu(x;\theta), \frac{\partial}{\partial\theta}\mu(x;\theta)\right), \ x \in \mathbb{R}. \end{aligned} \tag{6.102}$$

In this case S_{lin} is given by

$$S_{lin}(x) = \frac{1}{\sigma^2(x)} \begin{pmatrix} \frac{8-3\pi}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{8+\pi(3-8\log 2)}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{1}{-3+\log 16} \\ \frac{8+\pi(3-8\log 2)}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{8-3\pi}{4(-2+\pi\log 2)(-3+\log 16)} & \frac{1}{-3+\log 16} \\ \frac{1}{-3+\log 16} & \frac{1}{-3+\log 16} & \frac{-4+\log 16}{-3+\log 16} \end{pmatrix}, \ x \in \mathbb{R}, \tag{6.103}$$

and we find the following lower bound for the first term $\mathcal{V}_{-1,\theta}(\Delta, g, \hat{\theta})$ of the variance:

$$\begin{aligned} \mathcal{V}_{-1,\theta}(\Delta, g, \hat{\theta}) &= \Delta \frac{\mathbb{E}_{\nu,\theta}[Z_{lin}(\cdot)\, S_{lin}^{-1}(\cdot)\, Z_{lin}(\cdot)^T]}{\left(\mathbb{E}_{\nu,\theta}[Z_{lin}(\cdot)\Delta\, U_{lin}^{(1)}(\cdot)^T]\right)^2} \\ &\geq \frac{1}{\Delta}\left(\mathbb{E}_{\nu,\theta}\left[U_{lin}^{(1)}(\cdot)\, S_{lin}(\cdot)\, U_{lin}^{(1)}(\cdot)^T\right]\right)^{-1} \\ &= \frac{1}{\Delta}\left(\mathbb{E}_{\nu,\theta}\left[\frac{\frac{\partial}{\partial\theta}\mu(\cdot,\theta)}{\sigma^2(\cdot)}\right]\right)^{-1}. \end{aligned} \tag{6.104}$$

Again, the previous inequality holds on the premise that the expectations involved do not vanish. In this case an estimating function $g \in \mathcal{G}_\theta$ is small-Δ-optimal if there is a constant possibly depending on θ such that

$$Z_{lin} = K\, U_{lin}^{(1)} S_{lin}. \tag{6.105}$$

We stress that in our particular situation the vector $U_{lin}^{(1)} S_{lin}$ becomes

$$U_{lin}^{(1)}(\cdot) S_{lin}(\cdot) = \left(0,\, 0,\, \frac{\frac{\partial}{\partial\theta}\mu(\cdot,\theta)}{\sigma^2(\cdot)}\right). \tag{6.106}$$

This means that a small-Δ-optimal ordinary martingale estimating function is also small-Δ-optimal in the generalized model, and the lower bounds for the variance are the same. Compare also formula (6.20). When it comes to estimating the drift $\mu(\cdot,\theta)$, we cannot improve the model by incorporating the observations (H_Δ, L_Δ) to our analysis. We will try to give an explanation for this phenomenon below.

For the sake of completeness, let us briefly outline the situation for the linear estimating function, i.e. the martingale estimating function g which is defined by

$$g(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{3} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta), \tag{6.107}$$

with

$$\begin{aligned} k_1(\Delta, x, h, l, y; \theta) &= h - F^H(\Delta, x; \theta), \\ k_2(\Delta, x, h, l, y; \theta) &= l - F^L(\Delta, x; \theta), \\ k_3(\Delta, x, h, l, y; \theta) &= y - F^Y(\Delta, x; \theta), \end{aligned} \tag{6.108}$$

where

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[U], \quad \text{for } U \in \{H^Y_\Delta, L^Y_\Delta, Y_\Delta\}. \tag{6.109}$$

By the above statements, (6.107) is small-Δ-optimal if the weights (a_1, a_2, a_3) satisfy

$$\begin{aligned} a_1(\Delta, x; \theta) &= 0, \\ a_2(\Delta, x; \theta) &= 0, \\ a_3(\Delta, x; \theta) &= \frac{\partial}{\partial\theta}\mu(x,\theta) \Big/ \sigma(x)^2. \end{aligned} \tag{6.110}$$

Let us end this paragraph with some important remarks about generalized linear martingale estimating functions.

Assessment of the results for linear MEFs The small-Δ-asymptotic lower bound of the variance $\mathcal{V}_{-1,\theta}(\Delta, g, \hat{\theta})$, in a model of ordinary linear martingale estimating functions, is given by

$$\frac{1}{\Delta} \cdot \left(\mathbb{E}_{\nu,\theta} \left[\frac{\frac{\partial}{\partial\theta}\mu(\cdot,\theta)}{\sigma^2(\cdot)} \right] \right)^{-1}. \tag{6.111}$$

We obtain this result from our above considerations if (6.107) does not depend on h and l, that means if we neglect the observations H_Δ and L_Δ in our analyis. Note that we already encountered (6.111) when we were discussing ordinary small-Δ-optimal martingale estimating functions at the beginning of this chapter. Especially, compare (6.20). Formula (6.111) implies that the variance of an estimating function, that is constructed from a fixed number of equidistant observations n, explodes as $\Delta \to 0$. Or in other words, the approximately optimal linear estimator for the parameter θ of the drift $\mu(\cdot;\theta)$, inferred from the equidistant discrete sample $X_0, \ldots, X_T$, with $T = n\Delta$, is likely to perform badly if T is small. Incorporating the maximum and the minimum over the observation intervals does not improve the situation. The first term in the expansion of the asymptotic variance is the same. Compare formula (6.104). Usually, one would expect that additional data points result in a lower variance. Obviously, this is not the case here. The following deliberations give an explanation for this phenomenon. Let $(X_t = \mu t + B_t,\ 0 \le t \le 1)$ be a Brownian motion with drift. From Girsanov's Theorem we infer that the Log-likelihood is

$$\log L(\mu) = \mu X_1 - \frac{1}{2}\mu^2. \tag{6.112}$$

Hence, the maximum likelihood estimator for μ becomes

$$\hat{\mu}_{\mathrm{MLE}} = X_1. \tag{6.113}$$

Thus a sufficient statistic for the parameter μ is already given by X_1. In other words, in a Brownian model, all the information about the drift is contained in one single point, namely the endpoint X_1 of the trajectory $(X_t,\ 0 \le t \le 1)$. As we will see, there is no such flaw in the case of generalized estimating functions for $\sigma(\cdot;\theta)$.

A special quadratic estimator for the diffusion coefficient

In Proposition 6.3.2.5 we found a first order approximation of $\mathbb{E}_x[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, Y_\Delta)]$ of the form

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, Y_\Delta)\Big] = \Delta \mathbb{E}_{\nu,\theta}\Big[\mathcal{A}^S g_{0,\theta}(\cdot)\Big] + O(\Delta^{3/2}), \tag{6.114}$$

where the operator $\mathcal{A}^S$ is given by (6.72). If

$$\frac{\partial}{\partial h} g_{0,\theta}(x,h,x)\Big|_{h=x} = 0 \quad \text{and} \quad \frac{\partial}{\partial y} g_{0,\theta}(x,x,y)\Big|_{y=x} = 0, \tag{6.115}$$

a further expansion of $\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, Y_\Delta)\Big]$ is required. We found such an expansion for exactly this case in Proposition 6.3.2.6. Therefore we are equipped with the tools to state small-Δ-optimality for a class of martingale estimating functions that satisfy (6.115). For $x \in \mathbb{R}$, we define the vector $Z_{qua}(x)$ by

$$\begin{aligned} &Z_{qua}(x) \\ &= (Z_{qua,1}(x), Z_{qua,2}(x), Z_{qua,3}(x)) \\ &= \left(\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x)\Big|_{h=x}, \frac{\partial^2}{\partial y^2} g_{0,\theta}(x,x,y)\Big|_{y=x}, \frac{\partial}{\partial h}\frac{\partial}{\partial y} g_{0,\theta}(x,h,y)\Big|_{y=x} \right). \end{aligned} \tag{6.116}$$

Furthermore, we define U_{qua} to be the vector

$$U_{qua}(x) = \big(U_{qua,1}(x), U_{qua,2}(x), U_{qua,3}(x)\big),\ x \in \mathbb{R}, \tag{6.117}$$

having the entries

$$U_{qua,1}(x) = \left(\frac{2}{\pi} - 1\right)\theta\sigma(x)^2 \tag{6.118}$$

and

$$U_{qua,2}(x) = U_{qua,3}(x) = -\theta\sigma(x)^2. \tag{6.119}$$

Moreover, for $x \in \mathbb{R}$, let $S_{qua}^{-1}(x)$ be the matrix

$$S_{qua}^{-1}(x) = \theta^4\sigma^4(\cdot)\begin{pmatrix} \frac{1}{2} - \frac{4}{\pi^2} & \frac{1}{2}\left(\frac{1}{2} - \frac{2}{3\pi}\right) & \frac{1}{2}\left(\frac{7}{4} - \frac{4}{\pi}\right) \\ \frac{1}{2}\left(\frac{1}{2} - \frac{2}{3\pi}\right) & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}\left(\frac{7}{4} - \frac{4}{\pi}\right) & \frac{1}{2} & \frac{7}{4} - \frac{10}{3\pi} \end{pmatrix}, \tag{6.120}$$

whose inverse $S_{qua}(x)$ is given by

$$\frac{1}{\theta^4\sigma^4(x)}\begin{pmatrix} \frac{144\pi^2}{-608+168\pi+9\pi^2} & \frac{96\pi}{-608+168\pi+9\pi^2} & \frac{72\pi^2}{608-168\pi-9\pi^2} \\ \frac{96\pi}{-608+168\pi+9\pi^2} & \frac{18\left(2560-2112\pi+352\pi^2+21\pi^3\right)}{5(4864-3168\pi+432\pi^2+27\pi^3)} & -\frac{36\pi\left(-256+76\pi+3\pi^2\right)}{5(4864-3168\pi+432\pi^2+27\pi^3)} \\ \frac{72\pi^2}{608-168\pi-9\pi^2} & -\frac{36\pi\left(-256+76\pi+3\pi^2\right)}{5(4864-3168\pi+432\pi^2+27\pi^3)} & \frac{24\pi\left(-304+24\pi+27\pi^2\right)}{5(4864-3168\pi+432\pi^2+27\pi^3)} \end{pmatrix}. \tag{6.121}$$

With this notation we are able to write

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}(Y_0, H_\Delta^Y, Y_\Delta)\right] = \Delta\mathbb{E}_{\nu,\theta}[Z_{qua}(\cdot)U_{qua}(\cdot)^T] + O(\Delta^{3/2}) \tag{6.122}$$

and

$$\mathbb{E}_{\nu,\theta}\Big[g_{\Delta,\theta}^2(Y_0, H_\Delta^Y, Y_\Delta)\Big] = \Delta^2\mathbb{E}_{\nu,\theta}[Z_{qua}(\cdot)S_{qua}^{-1}(\cdot)Z_{qua}(\cdot)^T] + O(\Delta^{5/2}). \tag{6.123}$$

Thus, by means of Theorem 4.2.1.7, we find the following expansion of the variance of an estimator derived from a martingale estimating function g.

Theorem 6.3.3.2. *Suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 6.3.1.3. We assume that, for each $g \in \mathcal{G}_\theta$, the function*

$$(x, h, l, y) \longmapsto g_{\Delta,\theta}(x, h, l, y) \tag{6.124}$$

is independent of l, for all $\theta \in \Theta$ and for all $\Delta \geq 0$. Moreover, we assume that any $g \in \mathcal{G}_\theta$ satisfies (6.115)*. Let us assume that the coefficient μ of the underlying process*

X does not depend on θ. Then, for all $g \in \mathcal{G}_\theta$,

$$\mathrm{Var}_{\Delta,\nu,\theta}[g,\hat{\theta}] = \mathcal{V}_{0,\theta}(g,\hat{\theta}) + o(1), \tag{6.125}$$

where the first term on the right hand side of (6.125) *is lower bounded by*

$$\begin{aligned}\mathcal{V}_{0,\theta}(g,\hat{\theta}) &\geq \left(\mathbb{E}_{\mu,\theta}\left[U_{qua}(\cdot)\, S_{qua}(\cdot)\, U_{qua}(\cdot)^T\right]\right)^{-1} \\ &= \theta^2 \left(6 + \frac{96(28-9\pi)}{3\pi(56+3\pi)-608}\right)^{-1} \\ &\approx 0.33983 \cdot \theta^2, \end{aligned} \tag{6.126}$$

if the participating expectations do not vanish. Moreover, small-Δ-optimality holds for $g \in \mathcal{G}_\theta$ if there is a scalar $K \in \mathbb{R}$, possibly depending on θ, such that

$$Z_{qua} = K U_{qua} S_{qua}. \tag{6.127}$$

Proof. By definition of $\mathrm{Var}_{\Delta,\nu,\theta}[g,\hat{\theta}]$ in (6.86) one obtains the result by dividing (6.123) by formula (6.122) squared, in combination with the estimates in Lemma 6.2.0.13. □

Now, let g denote the particular quadratic martingale estimating function

$$g(\Delta,x,h,l,y;\theta) = \sum_{j=1}^{3} a_j(\Delta,x;\theta) k_j(\Delta,x,h,l,y;\theta) \tag{6.128}$$

with

$$\begin{aligned} k_1(\Delta,x,h,l,y;\theta) &= [h - F^H(\Delta,x;\theta)]^2 - \phi_{H,H}(\Delta,x;\theta), \\ k_2(\Delta,x,h,l,y;\theta) &= [y - F^Y(\Delta,x;\theta)]^2 - \phi_{X,X}(\Delta,x;\theta), \\ k_3(\Delta,x,h,l,y;\theta) &= [h - F^H(\Delta,x;\theta)][y - F^Y(\Delta,x;\theta)] - \phi_{H,X}(\Delta,x;\theta), \end{aligned} \tag{6.129}$$

where

$$F^U(\Delta,x;\theta) = \mathbb{E}_{x,\theta}[U], \quad \text{for } U \in \{Y_\Delta, H_\Delta^Y\}. \tag{6.130}$$

and

$$\phi_{U,V}(\Delta,x;\theta) = \mathrm{Cov}_{x,\theta}[U,V], \quad \text{for } U,V \in \{Y_\Delta, H_\Delta^Y\}. \tag{6.131}$$

According to our analysis, the small-Δ-optimal weights for the function (6.128) can be chosen as

$$\begin{aligned} a_1(\Delta,x;\theta) &= \frac{1}{2} \cdot 12.4711 \Big/ \sigma(x)^2, \\ a_2(\Delta,x;\theta) &= \frac{1}{2} \cdot 4.64644 \Big/ \sigma(x)^2, \end{aligned}$$

$$a_3(\Delta, x; \theta) = -6.23553 \Big/ \sigma(x)^2. \tag{6.132}$$

Assessment of the results for quadratic martingale estimating functions We stress that, if the martingale estimating function g does not depend on h, the lower bound for the first term $\mathcal{V}_{0,\theta}(g, \hat{\theta})$ in the expansion of the variance equals $0.5 \cdot \theta^2$. This result can easily be obtained as a special case of our analysis. But we already stated this fact in formula (6.24) in the introduction of this chapter. Note that $\sigma(x;\theta) = \theta\sigma(x)$ and thus $\left\{\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\right\}^2 = 4\theta^2\sigma(x)^4$. A comparison with formula (6.126) shows that we benefit from using generalized martingale estimating functions. In contrast to the case of linear martingale estimating functions for the drift parameter, incorporating the maximum or the minimum to the strictly quadratic model can significantly lower the variance of the resulting estimator of the diffusion parameter.

A range based estimator for the diffusion coefficient

In this section we finally want to consider martingale estimating functions that are independent of the variable y that corresponds to the end point Y_Δ. By the Propositions 6.3.2.3 and 6.3.2.5 we have the following expansions:

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y)\right] = \sqrt{\Delta}\,\mathbb{E}_{\nu,\theta}[Z_{range}(\cdot)U_{range}(\cdot)^T] \tag{6.133}$$

and

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H_\Delta^Y, L_\Delta^Y)\Big] = \Delta\,\mathbb{E}_{\nu,\theta}[Z_{range}(\cdot)S^{-1}_{range}(\cdot)Z_{range}(\cdot)^T] + o(\Delta), \tag{6.134}$$

where

$$Z_{range}(x) = \left(\frac{\partial}{\partial h} g_{0,\theta}(x,h,x)\Big|_{h=x}, \frac{\partial}{\partial l} g_{0,\theta}(x,x,l)\Big|_{l=x}\right), \; x \in \mathbb{R}, \tag{6.135}$$

and U_{range} is the vector

$$U_{range}(x) = (U^{(1/2)}_{lin,1}(x), U^{(1/2)}_{lin,2}(x)) = \left(\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}, -\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\right), \; x \in \mathbb{R}. \tag{6.136}$$

Note that the $U^{(1/2)}_{lin,i}$, $i = 1, 2$, would be zero if the diffusion coefficient of the underlying process X given by (6.9) did not not depend on θ. And moreover, the matrix S^{-1}_{range} is defined by

$$S^{-1}_{range}(x) = \theta^2\sigma^2(x)\begin{pmatrix} 1 - \frac{2}{\pi} & \frac{2}{\pi} + (1 - 2\log 2) \\ \frac{2}{\pi} + (1 - 2\log 2) & 1 - \frac{2}{\pi} \end{pmatrix}, \; x \in \mathbb{R}. \tag{6.137}$$

Its inverse is given by

$$S_{range} = \frac{1}{\theta^2\sigma^2(\cdot)} \begin{pmatrix} -\frac{\pi-2}{4(\log 2-1)(\pi\log 2-2)} & \frac{2+\pi-\pi\log 4}{4(\log 2-1)(\pi\log 2-2)} \\ \frac{2+\pi-\pi\log 4}{4(\log 2-1)(\pi\log 2-2)} & -\frac{\pi-2}{4(\log 2-1)(\pi\log 2-2)} \end{pmatrix}. \tag{6.138}$$

We find an expansion of the variance of a range based estimating function g_{range}. It is presented in the following theorem.

Theorem 6.3.3.3. *Suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 6.3.1.3. We assume that $\mathcal{G}_\theta$ is such that, for each $g \in \mathcal{G}_\theta$, the function*

$$(x,h,l,y) \longmapsto g_{\Delta,\theta}(x,h,l,y) \tag{6.139}$$

is independent of y, for all $\theta \in \Theta$ and for all $\Delta \geq 0$. If the coefficient σ of the underlying process X depends on θ, then, for all $g_{range} \in \mathcal{G}_\theta$,

$$\mathrm{Var}_{\Delta,\theta}[g_{range},\hat{\theta}] = \mathcal{V}_{0,\theta}(g_{range},\hat{\theta}) + o(1). \tag{6.140}$$

The first term in the expansion is lower bounded by

$$\begin{aligned} \mathcal{V}_{0,\theta}(g_{range},\hat{\theta}) &\geq \left(\mathbb{E}_{\nu,\theta}\left[U_{range}(\cdot)\,S_{range}(\cdot)\,U_{range}(\cdot)^T\right]\right)^{-1} \\ &= \theta^2\left(\frac{2}{\pi\log 2-2}\right)^{-1} \\ &\approx 0.088793\cdot\theta^2, \end{aligned} \tag{6.141}$$

provided that the participating expectations do not vanish. Finally, equality holds in (6.141) *and $g_{range} \in \mathcal{G}_\theta$ is small-Δ-optimal if there is a scalar $K \in \mathbb{R}$, possibly depending on θ, such that*

$$Z_{range} = KU_{range}S_{range}. \tag{6.142}$$

Proof. Recall the definition of $\mathrm{Var}_{\Delta,\nu,\theta}[g,\hat{\theta}]$ in (6.86). If one divides (6.134) by (6.133) squared, the result follows by means of Lemma 6.2.0.13. □

Let us state a very important fact. In order to prove Theorem 6.3.3.3 it is not necessary to work with the Lamperti transform of X and thus Assumption 6.3.1.1 is redundant. A first or second order expansion (with respect to $\sqrt{\Delta}$) of the expressions

$$\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(X_0,H_\Delta,L_\Delta)\right] \tag{6.143}$$

and

$$\mathbb{E}_{x,\theta}\left[g^2_{\Delta,\theta}(X_0,H_\Delta,L_\Delta)\right], \tag{6.144}$$

respectively, suffices to find asymptotic lower bounds for the variance of strictly range based martingale estimating functions. Consequently, we are able to work with any diffusion model (6.9) that satisfies the remaining assumptions of Section 6.3.1. Especially, we obtain the expansion

$$\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}(X_0,H_\Delta,L_\Delta)\right]=\sqrt{\Delta}\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}g_{0,\theta}(x,x,x)+O(\Delta), \tag{6.145}$$

with

$$\mathcal{A}_\theta^{(\frac{1}{2})}g(x,x,x)=\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\frac{\partial}{\partial h}g(x,h,x)\Big|_{h=x}-\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\frac{\partial}{\partial l}g(x,x,x)\Big|_{l=x}. \tag{6.146}$$

Compare also Remark 6.3.2.4 above. And the matrix S_{range} becomes

$$S_{range}=\frac{1}{\sigma^2(\cdot;\theta)}\begin{pmatrix}-\frac{\pi-2}{4(\log 2-1)(\pi\log 2-2)} & \frac{2+\pi-\pi\log 4}{4(\log 2-1)(\pi\log 2-2)}\\ \frac{2+\pi-\pi\log 4}{4(\log 2-1)(\pi\log 2-2)} & -\frac{\pi-2}{4(\log 2-1)(\pi\log 2-2)}\end{pmatrix}. \tag{6.147}$$

Therefore we find the lower bound

$$\begin{aligned}\mathcal{V}_{0,\theta}(g_{range},\hat{\theta})&\geq\left(\mathbb{E}_{\nu,\theta}\left[U_{range}(\cdot)\,S_{range}(\cdot)\,U_{range}(\cdot)^T\right]\right)^{-1}\\&=\left(\frac{2}{\pi\log 2-2}\right)^{-1}\left(\mathbb{E}_{\nu,\theta}\left[\frac{\left(\frac{\partial}{\partial\theta}\sigma(\cdot;\theta)\right)^2}{\sigma(\cdot;\theta)^2}\right]\right)^{-1}.\end{aligned} \tag{6.148}$$

To underpin its importance, let us state this result in a corollary.

Corollary 6.3.3.4. *We assume that we are given a stochastic differential equation*

$$dX_t=\mu(X_t;\theta)dt+\sigma(X_t;\theta)dB_t,\quad X_0=U,\ t\geq 0, \tag{6.149}$$

whose coefficients $\mu(\cdot;\theta)$ and $\sigma(\cdot;\theta)$ satisfy Assumption 6.2.0.8 and 6.2.0.9, but not necessarily Assumption 6.3.1.1. Let $\mathcal{G}_\theta$ be a class of flows that satisfies Assumption 6.3.1.3. We assume that $\mathcal{G}_\theta$ is such that, for each $g\in\mathcal{G}_\theta$, the function

$$(x,h,l,y)\longmapsto g_{\Delta,\theta}(x,h,l,y) \tag{6.150}$$

is independent of y for all $\theta\in\Theta$ and for all $\Delta\geq 0$. Set $v(\theta)=Q_\theta^\Delta(g_{\Delta,\theta}(\cdot)^2)$ and $\xi(\theta)=Q_\theta^\Delta(\partial_\theta g_{\Delta,\theta}(\cdot))$, where this time the probability measure Q_θ^Δ on $\mathbb{R}^3$ is defined by

$$Q_\theta^\Delta(x,h,l)=\nu_\theta(dx)\times f(\Delta,x,h,l;\theta), \tag{6.151}$$

The function $(h,l)\mapsto f(\Delta,x,h,l;\theta)$ denotes the joint density of the pair $(H_\Delta,L_\Delta)=\left(\sup_{0\leq t\leq\Delta}X_t,\inf_{0\leq t\leq\Delta}X_t\right)$, conditional on $X_0=x$. If the coefficient σ of the process

X depends on θ, then the variance

$$\operatorname{Var}_{\Delta,\nu,\theta}[g_{range}, \hat{\theta}] = \frac{v(\theta)}{\xi(\theta)^2} \tag{6.152}$$

satisfies, for all $g_{range} \in \mathcal{G}_\theta$, the expansion

$$\operatorname{Var}_{\Delta,\theta}[g_{range}, \hat{\theta}] = \mathcal{V}_{0,\theta}(g_{range}, \hat{\theta}) + o(1), \tag{6.153}$$

where the first term in the expansion is lower bounded by

$$\begin{aligned}
\mathcal{V}_{0,\theta}(g_{range}, \hat{\theta}) &\geq \left(\mathbb{E}_{\nu,\theta}\left[U_{range}(\cdot)\, S_{range}(\cdot)\, U_{range}(\cdot)^T\right]\right)^{-1} \\
&= \theta^2 \left(\frac{\pi \log 2}{2} - 1\right) \left(\mathbb{E}_{\nu,\theta}\left[\frac{\left(\frac{\partial}{\partial\theta}\sigma(\cdot;\theta)\right)^2}{\sigma(\cdot;\theta)^2}\right]\right)^{-1},
\end{aligned} \tag{6.154}$$

provided that the expectations involved do not vanish. Finally, equality holds in (6.154) *and $g_{range} \in \mathcal{G}_\theta$ is small-Δ-optimal if there is a scalar $K \in \mathbb{R}$, possibly depending on θ, such that*

$$Z_{range} = K U_{range} S_{range}. \tag{6.155}$$

Proof. The proof follows from formulae (6.145) and (6.146) in combination with Lemma 6.2.0.13. □

We end our analysis of the range based case by examining the concrete estimating function

$$g(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{2} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta) \tag{6.156}$$

with

$$\begin{aligned}
k_1(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)], \\
k_2(\Delta, x, h, l, y; \theta) &= [l - F^L(\Delta, x; \theta)],
\end{aligned}$$

where

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[U], \quad \text{for } U \in \{H_\Delta, L_\Delta\}. \tag{6.157}$$

Small-Δ-optimal weights for (6.156) are clearly given by

$$\begin{aligned}
a_1(\Delta, x; \theta) &= \sigma(x), \\
a_2(\Delta, x; \theta) &= -\sigma(x).
\end{aligned} \tag{6.158}$$

Assessment of the results for range based MEFs The factor $\left(\frac{2}{\pi \log 2 - 2}\right)^{-1} \approx 0.088793$ appearing in the asymptotic lower bounds of the variance for the class of range based estimating functions we displayed in formula (6.141) and formula (6.148) coincides with the variance of the expression

$$\sqrt{\frac{\pi}{2}} \frac{(H_\Delta^B - L_\Delta^B)}{\Delta}, \tag{6.159}$$

which is an unbiased estimator for the diffusion coefficient in the Brownian model without drift. As above H_Δ^B and L_Δ^B stand for $H_\Delta^B = \sup_{0 \le t \le \Delta} B_t$ and $L_\Delta^B = \inf_{0 \le t \le \Delta} B_t$, respectively. In order to see that the variances coincide, note again that $\mathbb{E}_0\left[H_\Delta^B \cdot L_\Delta^B\right] = \Delta(1 - 2\log 2)$. Consequently, we infer that

$$\begin{aligned}
\mathrm{Var}_0\left[\sqrt{\frac{2}{\pi}} \frac{(H_\Delta^B - L_\Delta^B)}{\Delta}\right] &= \frac{\pi}{2} \frac{\mathbb{E}_0[(H_\Delta^B)^2] + \mathbb{E}_0[(L_\Delta^B)^2] - 2\mathbb{E}_0\left[H_\Delta^B \cdot L_\Delta^B\right]}{\Delta} - 1 \\
&= \frac{\pi \log 2}{2} - 1.
\end{aligned} \tag{6.160}$$

Additionally, we can state that the range-based martingale estimating function is superior to the ordinary martingale estimating function in the following sense: in the statistical model, that is conceived for the Lamperti transform of a diffusion, the strictly range-based martingale estimating function has a small-Δ-asymptotic variance that is significantly lower than the one of the ordinary martingale estimating function. A comparison of formula (6.141), where we stated that the first term in the expansion of the variance is lower bounded by $0.088793 \cdot \theta^2$, with the lower bound for the ordinary martingale estimating function, which is given by $0.5 \cdot \theta^2$, shows that there is a gain in efficiency of about 82 %. The asymptotic lower bound for the variance in the strictly range-based model is even lower than the one for the variance of the generalized quadratic martingale estimating function, which is given by $0.33983 \cdot \theta^2$. See formula (6.126) and also the discussion in the paragraph "Assessment of the results for quadratic martingale estimating functions" above.

Eventually, let us stress again that, for the range-based estimating functions, we were even able to find a refinement, which is described in Corollary 6.3.3.4. Small-Δ-optimality of strictly range-based martingale estimating functions cannot only be stated in a model where the underlying process has the structure of a properly rescaled Lamperti transform, but for fairly general diffusions as well. Unfortunately, a comparison of this refined result with the corresponding result for ordinary martingale estimating function is not possible in general. Both lower bounds depend on $\sigma(\cdot;\theta)$ in a non-trivial way and the expressions are obviously not the same. Compare formula (6.154) and formula (6.24).

6.4 Case study

From now on, we consider an Ornstein-Uhlenbeck process defined by the following stochastic differential equation

$$dX_t = -\mu X_t dt + \sigma dB_t, \quad X_0 = x,\ t \geq 0. \tag{6.161}$$

Especially, we decided to choose $\mu \equiv 1$ and $\sigma \equiv 1$. Simulations were done for different values of Δ based on $n+1 = 501$ observations. Concretely, we considered $\Delta = 0.5$, 0.1, 0.01 and, for each value of Δ, 50 data sets were created. Therefore, our overall data set has the structure

$$X_0^{(j)}, H_\Delta^{(j)}, L_\Delta^{(j)}, X_\Delta^{(j)}, H_{2\Delta}^{(j)}, L_{2\Delta}^{(j)}, X_{2\Delta}^{(j)}, \ldots, H_T^{(j)}, L_T^{(j)}, X_T^{(j)}, \tag{6.162}$$

where

$$H_{i\Delta}^{(j)} = \sup_{(i-1)\Delta \leq s \leq i\Delta} X_s^{(j)}, \qquad L_{i\Delta}^{(j)} = \inf_{(i-1)\Delta \leq s \leq i\Delta} X_s^{(j)}, \tag{6.163}$$

for $i = 1, \ldots, n$, and with $T = n\Delta$, $\Delta = 0.5$, 0.1, 0.01, and $j = 1, \ldots, 50$. Recall that $n = 500$ is fixed. The 50 trajectories $\big(X_s^{(j)}, s \in [0, 250]\big)$, $j = 1, \ldots, 50$, from which the various samples are taken, are simulated according to an Euler-scheme. Different segments of the trajectories are simulated with different accuracy. This means that the first interval $[0, 5]$ contains 500000 simulated equidistant values of X. The time interval [5,50] contains 450000 simulated equidistant values of X, and the final time interval [50,250] contains 400000 simulated equidistant values of X. Consequently, there are at least 1000 simulated values of X in each observation interval $((i-1)\Delta, i\Delta]$, for $i = 1, \ldots, 500$ and for every value of $\Delta = 0.5$, 0.1, 0.01.

In the first paragraph we assume that $\sigma = 1$ is known and we consider estimators of μ, whereas in the second paragraph we assume that the parameter $\mu = 1$ is known and we consider estimators of σ^2. The mean value and the standard deviation of the different estimators are given in the tables of the respective paragraphs. Moreover, the columns labeled "min" and "max" indicate the range of the 50 estimators we calculated.

6.4.1 Drift estimation for an OU-process

Here, we assume that $\sigma \equiv 1$ is fixed and we consider the ordinary martingale estimating function given by

$$g_{lin,ord}(\Delta, x, h, l, y; \theta) = a_1(\Delta, x; \theta) k_1(\Delta, x, h, l, y; \theta), \tag{6.164}$$

where

$$k_1(\Delta, x, h, l, y; \theta) = [y - F^X(\Delta, x; \theta)], \tag{6.165}$$

with

$$F^X(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[X_\Delta], \quad \text{for } U \in \{H_\Delta, L_\Delta, X_\Delta\}. \tag{6.166}$$

We approximate the above expectation by

$$\mathbb{E}_{x,\theta}[X_\Delta] \approx x - \mu x \Delta. \tag{6.167}$$

Note that the resulting estimators are biased, due to these approximations. Moreover, the small-Δ-optimal weight for the linear martingale estimating function $g_{lin,ord}$ can be chosen as

$$a_1(\Delta, x; \theta) = x. \tag{6.168}$$

Note that this estimating function is also the small-Δ-optimal martingale estimating function in the generalized model. An explanation for this phenomenon was given in the paragraph "Assessment of the results for linear MEFs" in Section 6.3.3. Consequently, the ordinary martingale estimating function is definitely the first choice when it comes to estimating the drift coefficient, and inference in this case should be made by means of estimating functions that are designed for equidistant observations alone.

Solving the estimating equation results in the following table:

Table 6.1: Estimators for the drift μ

Version	Δ	mean	std. dev.	max	min
Ordinary linear MEF	0.5	0.80	0.072	1.016	0.679
	0.1	1.02	0.233	1.79	0.612
	0.01	1.439	1.032	5.141	0.419

Recall that the estimators are inferred from samples that are generated with drift parameter $\mu = 1$ and that the sample size $n = 500$ is fixed. In agreement with the theoretical results, the estimator of μ deteriorates when Δ becomes small.

6.4.2 Estimation of the diffusion coefficient in an OU-model

We consider different types of martingale estimating functions for the diffusion coefficient. First, we compare the estimator inferred from the ordinary martingale estimating function to the one inferred from a generalized martingale estimating function. Then, we compare the ordinary estimator to a range based estimator.

Ordinary vs. Generalized Martingale Estimating Functions

We assume that $\mu \equiv 1$ is fixed and we compare two different estimators of σ^2. Namely, we consider the ordinary martingale estimating function

$$g_{qua,ord}(\Delta, x, h, l, y; \theta) = a_2(\Delta, x; \theta) k_2(\Delta, x, h, l, y; \theta) \tag{6.169}$$

and compare it to

$$g_{qua,gen}(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{3} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta), \tag{6.170}$$

where

$$\begin{aligned}
k_1(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)]^2 - \phi_{H,H}(\Delta, x; \theta), \\
k_2(\Delta, x, h, l, y; \theta) &= [y - F^Y(\Delta, x; \theta)]^2 - \phi_{X,X}(\Delta, x; \theta), \\
k_3(\Delta, x, h, l, y; \theta) &= [h - F^H(\Delta, x; \theta)][y - F^X(\Delta, x; \theta)] - \phi_{H,X}(\Delta, x; \theta),
\end{aligned} \tag{6.171}$$

with

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[U], \quad \text{for } U \in \{H_\Delta, X_\Delta\}, \tag{6.172}$$

and

$$\phi_{U,V}(\Delta, x; \theta) = \mathrm{Cov}_{x,\theta}[U, V], \quad \text{for } U, V \in \{H_\Delta, X_\Delta\}. \tag{6.173}$$

The above expressions are approximated by

$$\begin{aligned}
\left(y - \mathbb{E}_{x,\theta}[X_\Delta]\right)^2 - \mathrm{Var}_{x,\theta}[X_\Delta] = y^2 &- 2y(x + \mu(x)\Delta) + 2(x^2 + 2x\mu(x)\Delta) \\
&- (x^2 + 2x\mu(x)\Delta + \sigma^2\Delta) + O(\Delta^2),
\end{aligned} \tag{6.174}$$

$$\begin{aligned}
\left(h - \mathbb{E}_{x,\theta}[H_\Delta]\right)^2 - \mathrm{Var}_{x,\theta}[H_\Delta] = h^2 &- 2h\left(x + \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{1}{2}\mu(x)\Delta\right) \\
&+ 2\left(x^2 + \frac{2}{\pi}\sigma^2\Delta + 2x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + x\mu(x)\Delta\right) \\
&- \left(x^2 + 2x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + x\mu(x)\Delta + \sigma^2\Delta\right) \\
&+ O(\Delta^{3/2}),
\end{aligned} \tag{6.175}$$

and

$$\begin{aligned}
&\left(h - \mathbb{E}_{x,\theta}[H_\Delta]\right)\left(y - \mathbb{E}_{x,\theta}[X_\Delta]\right) - \mathrm{Cov}_{x,\theta}[H_\Delta, X_\Delta] \\
&\quad = hy - y\left(x + \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{1}{2}\mu(x)\Delta\right) - 2h(x + \mu(x)\Delta) \\
&\quad + 2\left(x^2 + x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{3}{2}\mu(x)\Delta\right) \\
&\quad - \left(x^2 + x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{3}{2}\mu(x)\Delta + \frac{1}{2}\sigma^2\Delta\right) + O(\Delta^{3/2}).
\end{aligned} \tag{6.176}$$

The small-Δ optimal weights for the generalized martingale estimating function $g_{qua,soph}$ can be chosen as

$$\begin{aligned} a_1(\Delta, x; \theta) &= \frac{1}{2} \cdot 12.4711, \\ a_2(\Delta, x; \theta) &= \frac{1}{2} \cdot 4.64644, \\ a_3(\Delta, x; \theta) &= -6.23553. \end{aligned} \tag{6.177}$$

We obtain the following table for the estimators of the true parameter $\sigma^2 = 1$.

Table 6.2: Quadratic estimators for the diffusion parameter σ^2

Version	Δ	mean	std. dev.	max	min
Ordinary MEF	0.5	0.395	0.025	0.434	0.328
	0.1	0.862	0.065	1.031	0.735
	0.01	0.99	0.074	1.188	0.843
Generalized MEF	0.5	0.686	0.037	0.759	0.605
	0.1	0.928	0.06	1.099	0.808
	0.01	1.005	0.068	1.20	0.867

Comparing the lines belonging to $\Delta = 0.01$ in the previous table, we see that the estimator $\hat{\sigma}^2_{qua,gen}$ inferred from the generalized martingale estimating function has slightly smaller bias and significantly smaller standard deviation than the estimator $\hat{\sigma}^2_{qua,ord}$ inferred from the ordinary martingale estimating function. Let us compare the mean squared errors of $\hat{\sigma}^2_{qua,gen}$ and $\hat{\sigma}^2_{qua,ord}$ for $\Delta = 0.01$. We have

$$\mathbb{E}[(\hat{\sigma}^2_{qua,gen} - \sigma^2)^2] = 0.005^2 + 0.068^2 = 0.004649 \tag{6.178}$$

and

$$\mathbb{E}[(\hat{\sigma}^2_{qua,ord} - \sigma^2)^2] = 0.01^2 + 0.074^2 = 0.005576. \tag{6.179}$$

The quotient of both quantities is

$$\frac{\mathbb{E}[(\hat{\sigma}^2_{qua,gen} - \sigma^2)^2]}{\mathbb{E}[(\hat{\sigma}^2_{qua,ord} - \sigma^2)^2]} = 0.833752. \tag{6.180}$$

Evidently, the generalized martingale estimating function is superior to the ordinary martingale estimating function, even though we do not exactly discover the gain in efficiency that was predicted by our theoretical results, see the paragraph "Assessment of the results for quadratic martingale estimating functions" above. This mismatch between theory and practice might be due to the discretization of the trajectories. For more details about the discretization error, see also the upcoming discussion for the range based martingale estimating functions.

Ordinary vs. Range Based Martingale Estimating Functions

Again, we assume that $\mu \equiv 1$ is known. We want to estimate σ^2. Here, we compare the ordinary quadratic martingale estimating function $g_{qua,ord}$ given by (6.169) to the range based martingale estimating function

$$g_{range}(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{2} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta), \tag{6.181}$$

where

$$\begin{aligned} k_1(\Delta, x, h, l, y; \theta) &= h - F^H(\Delta, x; \theta), \\ k_2(\Delta, x, h, l, y; \theta) &= l - F^L(\Delta, x; \theta), \end{aligned}$$

with

$$F^U(\Delta, x; \theta) = \mathbb{E}_{x,\theta}[U], \quad \text{for } U \in \{H_\Delta, L_\Delta\}. \tag{6.182}$$

We consider the following approximations:

$$h - \mathbb{E}_{x,\theta}[H_\Delta] = h - x - \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{1}{2}\mu x \Delta + O(\Delta^{3/2}) \tag{6.183}$$

and

$$l - \mathbb{E}_{x,\theta}[L_\Delta] = l - x + \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{1}{2}\mu x \Delta + O(\Delta^{3/2}). \tag{6.184}$$

Moreover, the small-Δ-optimal weights can be chosen as

$$\begin{aligned} a_1(\Delta, x; \theta) &= 1, \\ a_2(\Delta, x; \theta) &= -1. \end{aligned} \tag{6.185}$$

The results for the range based estimators for the true parameter $\sigma^2 = 1$ for different values of Δ are displayed in Table 6.3. We compare them to the estimators inferred from the ordinary martingale estimating function.

Table 6.3: Range-based estimators for the diffusion parameter σ^2

Version	Δ	mean	std. dev.	max	min
Ordinary MEF	0.5	0.395	0.025	0.434	0.328
	0.1	0.863	0.065	1.031	0.735
	0.01	0.99	0.074	1.188	0.843
Range Based MEF	0.5	0.908	0.024	0.961	0.861
	0.1	0.945	0.032	1.039	0.868
	0.01	0.958	0.032	1.05	0.879

It seems that the range based estimation method does not work properly for small Δ. For $\Delta = 0.01$ the corresponding estimators take a value of approximately 0.96 instead of 1. This contradiction can be explained by the discretization. Let $(B_t,\ 0 \le t \le 1)$ be a standard Brownian motion of $\mathbb{R}$ and let $(B_t^{(sim)},\ 0 \le t \le 1)$ denote a simulation of B that takes $\mathcal{N}$ equidistant time steps on $[0,1]$. Rogers and Satchell [61] proved that

$$H_1 = H_1^{(sim)} + \delta, \tag{6.186}$$

where $H_1 = \sup_{0\le t\le 1} B_t$, $H_1^{(sim)} = \sup_{0\le t\le 1} B_t^{(sim)}$ and δ is a random variable that satisfies

$$\mathbb{E}_0[\delta] = \sqrt{2\pi}\left(\frac{1}{4} - \frac{(\sqrt{2}-1)}{6}\right)\sqrt{\frac{1}{\mathcal{N}}} \tag{6.187}$$

and

$$\mathbb{E}_0[\delta^2] = \frac{\left(1+\frac{3\pi}{4}\right)}{12}\frac{1}{\mathcal{N}}. \tag{6.188}$$

In other words the maximum is underestimated by a certain amount δ with mean (6.187). In our case this implies the following. For small values of Δ, we are approximately in a Brownian setting and thus the estimator $\hat{\sigma}_{range}$ derived from the first order expansions of g_{range} satisfies

$$\mathbb{E}_x\left[\hat{\sigma}_{range}\right] = \frac{\mathbb{E}_x[H_\Delta - L_\Delta - 2\delta]}{2\sqrt{\frac{2}{\pi}}\sqrt{\Delta}} \approx 1 - \frac{\pi\left(\frac{1}{4} - \frac{(\sqrt{2}-1)}{6}\right)}{\sqrt{\mathcal{N}}}. \tag{6.189}$$

The trajectories of the Ornstein-Uhlenbeck process we used for our estimation were simulated in such a way that there were at least $\mathcal{N} = 1000$ equidistant simulated values of X in each observation interval – even for the smallest value of $\Delta = 0.01$. For $\mathcal{N} = 1000$ the right hand side of (6.189) equals 0.982. Hence, a range based estimator $\hat{\sigma}_{range}$ inferred from g_{range} is rather an estimator of 0.982 than of 1. Finally, for $\mathcal{N} = 1000$, formula (6.189) implies that

$$\left(\mathbb{E}_x\left[\hat{\sigma}_{range}\right]\right)^2 \approx 0.964. \tag{6.190}$$

This is approximately the value we discovered in Table 6.3 for the range based estimator when $\Delta = 0.01$. It is straightforward to calculate the adjusted range based estimators. Their figures are displayed in Table 6.4.

As we expected, for small values of Δ, the range based estimator has smaller variance than the estimator inferred from the ordinary quadratic estimating function $g_{qua,ord}$. It is also superior to the estimator inferred from the generalized quadratic martingale estimating function $g_{qua,gen}$. Even for the not adjusted range-based estimator, this effect is clearly visible. Compare the lines belonging to $\Delta = 0.01$ in Table 6.3 and Table 6.4. Also see Table 6.2 in the previous paragraph. Concretely, a comparison of the biases

Table 6.4: Adjusted range-based estimators for σ^2

Version	Δ	mean	std. dev.	max	min
Ordinary MEF	0.5	0.395	0.025	0.434	0.328
	0.1	0.863	0.065	1.031	0.735
	0.01	0.99	0.074	1.188	0.843
Adjusted Range Based MEF	0.5	0.943	0.025	0.997	0.895
	0.1	0.98	0.032	1.076	0.902
	0.01	0.993	0.033	1.088	0.913

and the standard deviations for $\Delta = 0.01$ in Table 6.4 shows that

$$\mathbb{E}[(\hat{\sigma}^2_{qua,ord} - \sigma^2)^2] = 0.01^2 + 0.074^2 = 0.00547 \tag{6.191}$$

and

$$\mathbb{E}[(\hat{\sigma}^2_{range} - \sigma^2)^2] = 0.007^2 + 0.033^2 = 0.00114. \tag{6.192}$$

The quotient of both quantities is

$$\frac{\mathbb{E}[(\hat{\sigma}^2_{range} - \sigma^2)^2]}{\mathbb{E}[(\hat{\sigma}^2_{qua,ord} - \sigma^2)^2]} = 0.20778. \tag{6.193}$$

This means that the mean squared error for the range-based model is about 80 % lower than the mean squared error for the ordinary estimating function. This almost corresponds to the theoretical values we discovered in Section 6.3.3. Especially, see the paragraph "Assessment of the results for range based MEFs" therein.

It would be interesting to know if this effect carries over to martingale estimating function constructed with triplets of observations $(H_\Delta, L_\Delta, X_\Delta)$. The question is whether the lower bound of the variance of such an estimating function is even smaller than the one of the range-based martingale estimating function obtained from the pair (H_Δ, L_Δ). We have cause to believe that this is the case. But a concise proof remains to be conducted.

6.4.3 Annotations

In our simulation study we did not only replace the optimal weights $a_1(\Delta, x; \theta)$, $a_2(\Delta, x; \theta)$ and $a_3(\Delta, x; \theta)$ by the respective small-Δ-optimal weights, but we also approximated the expectations

$$\mathbb{E}_{x,\theta}[H_\Delta], \quad \mathbb{E}_{x,\theta}[L_\Delta] \quad \text{and} \quad \mathbb{E}_{x,\theta}[X_\Delta], \tag{6.194}$$

and the respective variances by their second order approximations with respect to $\sqrt{\Delta}$. As we saw above, the resulting estimators were biased. This is due to the fact that the martingale property of the estimating functions is destroyed by the additional approximation. We deliberately ignored that the behavior of such estimating functions and the asymptotics of the resulting estimators are not covered by the theoretical results we proved so far. However, the correctness of the simulations justifies their use in hindsight. Of course this reasoning is not a substitute for a concise proof. But note that there are results available for exactly this situation for ordinary martingale estimating functions.

Recall the small-Δ-optimal ordinary martingale estimating functions given by formula (6.6) and formula (6.8) in the introduction of this chapter. By replacing the expectation $\mathbb{E}_{x,\theta}[X_\Delta]$ and the variance $\mathrm{Var}_{x,\theta}[X_\Delta]$ by their first order approximations with respect to Δ, one obtains the approximated small-Δ-optimal estimating functions

$$\sum_{i=1}^{n} \frac{\partial_\theta \mu(X_{t_{i-1}};\theta)}{\sigma^2(X_{t_{i-1}};\theta)}[X_{t_i} - X_{t_{i-1}} - \Delta_i \mu(X_{t_{i-1}};\theta)] \tag{6.195}$$

and

$$\begin{aligned}&\sum_{i=1}^{n} \frac{\partial_\theta \mu(X_{t_{i-1}};\theta)}{\sigma^2(X_{t_{i-1}};\theta)}[X_{t_i} - X_{t_{i-1}} - \Delta_i \mu(X_{t_{i-1}};\theta)] \\ &\quad + \sum_{i=1}^{n} \frac{\partial_\theta \sigma^2(X_{t_{i-1}};\theta)}{2\sigma^4(X_{t_{i-1}};\theta)\Delta_i}[(X_{t_i} - X_{t_{i-1}} - \Delta_i \mu(X_{t_{i-1}};\theta))^2 - \Delta_i \sigma^2(X_{t_{i-1}};\theta)],\end{aligned} \tag{6.196}$$

respectively. The approximate estimating functions (6.195) and (6.196) result in biased estimators. The bias can be serious if Δ is not small. However, if Δ is sufficiently small the inferred estimators might work all right. Consistency is proved in the article of Florens-Zmirou [26] when the asymptotics satisfy $\Delta_n \to 0$ and $n\Delta_n \to \infty$. Moreover, asymptotic normality can be stated on the further condition $n\Delta_n^3 \to 0$. Related results were also proved by Yoshida [72]. And finally, Kessler [44] used higher order expansions of the moments of the transition distribution to obtain estimators that are asymptotically normal, even when Δ_n tends more slowly to zero as n tends to ∞.

A result, analogous to the one of Florens-Zmirou, for our generalized martingale estimating functions will not be derived. This goes beyond the scope of this thesis. But of course this issue gives rise to further research.

7 Extensions

7.1 Introduction & Motivation

In Chapter 5 we found an expansion of the expression $\mathbb{E}_x[g(H_t^Y, L_t^Y, Y_t)]$ with respect to $\sqrt{t}$ including powers of 2, where $g : \mathbb{R}^3 \to \mathbb{R}$ can be any sufficiently smooth function that does not grow too fast. Recall that Y denoted the Lamperti transform of $(X_t,\, t \geq 0)$. The process X in turn was assumed to be a diffusion that satisfies a stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x,\ t \geq 0, \tag{7.1}$$

with sufficiently smooth coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$. Moreover, the processes H^Y and L^Y were defined by $H_t^Y = \sup_{0 \leq s \leq t} Y_s$ and $L_t^Y = \inf_{0 \leq s \leq t} Y_s$. The reason why we considered the Lamperti transform of X was a technical one. For a diffusion process with non-constant diffusion coefficient it is only possible to calculate the first coefficient of the expansion with respect to $\sqrt{t}$. But even for the Lamperti transform, the techniques of Chapter 5 do not permit to calculate coefficients of any order higher than 2. Nevertheless, our expansions of the expression $\mathbb{E}_x[g(H_t^Y, L_t^Y, Y_t)]$ were sufficient to establish a theory for approximately optimal martingale estimating functions on small observation intervals in a model where the drift $\mu(\cdot;\theta)$ and the diffusion coefficient $\sigma(\cdot;\theta)$ of X are parameterized by a parameter $\theta \in \Theta \subset \mathbb{R}$. Since the length of the observation intervals was denoted with Δ instead of t, we called the resulting martingale estimating functions *small-Δ-optimal martingale estimating functions*. The results can be found in Chapter 6. In Section 6.4, a simulation study for an Ornstein-Uhlenbeck process showed that for small values of Δ the estimators inferred from approximations of the optimal estimating functions yield reasonable results. But we must also take into consideration that these estimators were flawed inasmuch as they were significantly biased for large values of Δ. It would be desirable to have an overall expansion of $\mathbb{E}_x[g(H_\Delta, L_\Delta, X_\Delta)]$ with respect to $\sqrt{\Delta}$, for in that case more accurate estimators could be determined.

The above remarks give a motivation to find an overall expansion of the joint probability of (H_t, X_t) with respect to $\sqrt{t}$. This is the aim of the present chapter. The vector (H_t, L_t, X_t) is much harder to deal with. We entirely neglect the problem of determining the joint probability of the triplet (H_t, L_t, X_t). The goal of expanding the joint probability of (H_t, X_t) will be achieved via a detour. We analyze the behavior of the first hitting time of a pinned diffusion first. In other words, we focus our attention on the expression

$$\mathbb{P}_x\big[\tau_h \le t \,|\, X_t = y\big], \tag{7.2}$$

where $\tau_h = \inf\{t > 0 \,|\, X_t \ge h\}$ and $x, y < h$. This quantity and similar ones have already been studied by several authors. An important reference is the article of Borovkov and Downes [17] in which a relationship between the asymptotic form of conditional boundary crossing probabilities and first-passage time densities of diffusion processes is established. However, more relevant to our studies is the work of Baldi and Caramellino [6]. We quote a special case of their results in Section 7.2, see Proposition 7.2.1.2, in particular. The contribution of Baldi and Caramellino is a description of the exact asymptotics of the hitting probabilities of a diffusion X, pinned at $X_t = y$, as t tends to 0. Their approach mainly relies on large deviation techniques and results in a first order expansion of (7.2). We are going to extend their findings and give an overall expansion of (7.2) for a certain class of diffusion processes X with respect to t. Once the expansion of the quantity (7.2) is determined, an expansion of the transition density $p(t, x, y)$ can be used to approximate the joint probability of (H_t, X_t). The function p can be expanded in different ways. An important example is the so-called WKB-expansion found by Kampen [42]. Another expansion of the transition density was derived by Yacine Aït-Sahalia, see [3] or [4]. His approach yields a complete expansion of $p(t, x, y)$ with respect to $\sqrt{t}$. In order to give a concrete example, we will use our results to calculate a fourth order expansion of $\mathbb{E}_x[g(H_t, X_t)]$ in Chapter 8. In order to determine this expansion, we will make use of Yacine Aït-Sahalia's results, since his representation of $p(t, x, y)$ is most convenient for our purposes.

In the present chapter, we proceed as follows. First, in Section 7.2, we recite some very important results of Baldi and Caramellino in detail and adjust them to our purposes. The most interesting thing about this section is a transformation that reduces the problem of determining an expansion of (7.2) to the problem of calculating an expansion of a functional of a rescaled Brownian bridge process. Therefore, in Section 7.3, we analyze this special functional of the Brownian bridge. An overall expansion of the respective functional, with respect to the scaling parameter, can be derived by means of elaborate partial differential equation techniques. In Section 7.4 we state an important convergence result for the series expansions of Section 7.3. And finally, a retransformation of the expansions derived in Section 7.3 results in an overall expansion of the expression (7.2). This final series expansion is calculated and displayed in Section 7.5.

7.2 Some large deviation techniques and their implications

Formula (5.84), which can be considered as the core result of Section 5.2, is based on elementary estimates of diffusion processes and their sample paths. However, the applied techniques do not allow a generalization of this result. In order to determine higher order terms of the expansion, some auxiliary results are required. We present different basic facts in the sequel. Moreover, we indicate some interesting implications of these facts.

7.2.1 Asymptotical behaviour of one-dimensional pinned diffusions

In this paragraph we present a result that Baldi and Caramellino proved in their article [6]. We keep our notations very close to the notations in this article.

Let us consider a general one-dimensional, time-homogeneous diffusion process X on $\mathbb{R}$, i.e. a process which satisfies the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x, \ t \geq 0, \tag{7.3}$$

with a sufficiently smooth drift coefficient μ. In order to simplify our upcoming analysis, we assume that the diffusion coefficient σ is constant.

Assumption 7.2.1.1. *The diffusion coefficient σ of the process X in* (7.3) *constantly equals 1.*

Note that this is no restriction. By means of the Lamperti transform, a diffusion process X can be transformed into a diffusion Y with constant diffusion coefficient 1. The Lamperti transform is given by $Y_t = F(X_t)$, where F can be any primitive of $1/\sigma$. By Itô's formula, the process Y satisfies the stochastic differential equation

$$dY_t = \left(\frac{\mu(F^{-1}(Y_t))}{\sigma(F^{-1}(Y_t))} - \frac{1}{2}\sigma'(F^{-1}(Y_t)) \right) dt + dB_t, \quad Y_0 = \xi, \ t \geq 0, \tag{7.4}$$

where $\xi = F(x)$. If σ is uniformly elliptic, the probability of X, pinned at $X_t = y$, to cross level h equals the probability of Y, pinned at $Y_t = \eta = F(y)$, to cross level $\mathbb{h} = F(h)$. Therefore it is possible to translate the results for the case of a constant diffusion coefficient $\sigma > 0$ to a general setting where $\sigma : \mathbb{R} \to \mathbb{R}_+$ is a suitable function.

Moreover, throughout this section we implicitly assume that X has a transition density. Sufficient conditions for this to hold can be found in Karatzas and Shreve [43] for example. Already, the results we found in Chapter 3 imply the existence of a transition density.

Let $l \leq x, y \leq h$ and let X be a diffusion process starting in x. We want to consider the first hitting times

$$\tau_h = \inf \{ t > 0 \,|\, X_t \geq h \} \tag{7.5}$$

and

$$\tau_l = \inf \{ t > 0 \,|\, X_t \leq l \}. \tag{7.6}$$

Moreover, let $\tau_{[l,h]}$ denote the first exit time of X from the interval (l, h), i.e.

$$\tau_{[l,h]} = \inf \{ t > 0 \,|\, X_t \notin (l, h) \}. \tag{7.7}$$

The aim of the present section is to analyze these stopping times of X, conditional on

$X_t = y$, and to describe their asymptotical behavior as $t \to 0$. Concretely, we want to find the exact asymptotics of the one barrier problems

$$\mathbb{P}_x\big[\tau_h \leq t \big| X_t = y\big] \quad \text{and} \quad \mathbb{P}_x\big[\tau_l \leq t \big| X_t = y\big], \tag{7.8}$$

and the two barrier problem

$$\mathbb{P}_x\big[\tau_{[l,h]} \leq t \big| X_t = y\big], \tag{7.9}$$

as t tends to 0. In the sequel, we will sometimes refer to x and to y as the starting point and the pinning point, respectively. In the next section we are going to regard the above probabilities as functions of (t, x, h, l, y). But for now, let us assume that the pinning point y is fixed.

We modify our notations slightly. Let $\epsilon > 0$. In order to normalize the problem in some sense, we make a switch in the time variable and set $U_t^\epsilon = X_{\epsilon t}$. The process U^ϵ solves the stochastic differential equation

$$dU_t^\epsilon = \epsilon\, \mu(U_t^\epsilon)dt + \sqrt{\epsilon} dB_t, \quad U_0^\epsilon = x,\ 0 \leq t \leq 1. \tag{7.10}$$

Moreover, we define the process

$$W_x^\epsilon(t) = x + \sqrt{\epsilon} B_t, \quad 0 \leq t \leq 1. \tag{7.11}$$

Note that this process would be the same as U^ϵ if $\mu \equiv 0$.

As usual, we associate processes with measures on the space of continuous sample paths $\mathcal{C} = \mathcal{C}([0,1], \mathbb{R})$. On $\mathcal{C}$ the coordinate variable process X is defined by $X_t(\omega) = \omega_t$. Moreover, from X we can infer the filtration $(\mathcal{F}_t,\, 0 \leq t \leq 1)$. It is given by $\mathcal{F}_t = \sigma(X_s,\, 0 \leq s \leq t)$. Apart from the Markov measure $\mathbb{P}_x$ making $\mathbb{P}_x[X_0 = x] = 1$ and the associated expectation operator $\mathbb{E}_x$, we consider the following four probability laws on $(\mathcal{C}, \mathcal{F}_1)$:

$$\begin{aligned} \mathbb{P}_x^\epsilon &= \text{ the law of } W_x^\epsilon, \\ \hat{\mathbb{P}}_x^{y,\epsilon} &= \text{ the law of } W_x^\epsilon \text{ pinned by } W_x^\epsilon(1) = y, \\ \mathbb{Q}_x^\epsilon &= \text{ the law of } U_x^\epsilon, \\ \hat{\mathbb{Q}}_x^{y,\epsilon} &= \text{ the law of } U_x^\epsilon \text{ pinned by } U_x^\epsilon(1) = y. \end{aligned} \tag{7.12}$$

By definition it makes no difference if we examine the hitting probabilities of the conditioned process $\hat{U}^{y,\epsilon}$, pinned by $U_1^\epsilon = y$, or if we examine the hitting probabilities of the diffusion process X, pinned by $X_\epsilon = y$. In other words

$$\mathbb{P}_x\big[\tau_h \leq \epsilon \big| X_\epsilon = y\big] = \hat{\mathbb{Q}}_x^{y,\epsilon}[\tau_h \leq 1]. \tag{7.13}$$

Analogously, we obtain the relations

$$\mathbb{P}_x\big[\tau_l \leq \epsilon \big| X_\epsilon = y\big] = \hat{\mathbb{Q}}_x^{y,\epsilon}[\tau_l \leq 1] \tag{7.14}$$

and

$$\mathbb{P}_x\big[\tau_{[l,h]} \leq \epsilon \big| X_\epsilon = y\big] = \hat{\mathbb{Q}}_x^{y,\epsilon}[\tau_{[l,h]} \leq 1]. \tag{7.15}$$

The main idea is to show that $\hat{\mathbb{Q}}_x^{y,\epsilon}$ has a density with respect to $\hat{\mathbb{P}}_x^{y,\epsilon}$ and to calculate this density approximately. If

$$\zeta_t = \zeta_t(\epsilon) = \exp\left(\int_0^t \mu(X_u)dX_u - \frac{\epsilon}{2}\int_0^t \mu(X_u)^2 du\right), \tag{7.16}$$

then by Girsanov's theorem, for every $A \in \mathcal{F}_1$,

$$\mathbb{Q}_x^\epsilon[A] = \mathbb{E}_x^\epsilon\big[\zeta_1 \mathbb{1}_{\{X \in A\}}\big]. \tag{7.17}$$

By the fact that U_x^ϵ and W_x^ϵ are random variables that are absolutely continuous with respect to the Lebesgue measure and by Lemma 3.1 in [6] it is possible to derive the following relation between $\hat{\mathbb{P}}_x^{y,\epsilon}$ and $\hat{\mathbb{Q}}_x^{y,\epsilon}$

$$\hat{\mathbb{Q}}_x^{y,\epsilon}[A] = \frac{p_\epsilon(1,x,y)}{q_\epsilon(1,x,y)}\hat{\mathbb{E}}_x^{y,\epsilon}[\zeta_1 \mathbb{1}_{\{X\in A\}}], \quad \forall A \in \mathcal{F}_1, \tag{7.18}$$

where $p_\epsilon(t,x,y)$ and $q_\epsilon(t,x,y)$ denote the transition densities of W_x^ϵ and U_x^ϵ, respectively.

Let G denote a primitive of μ. This means $G(y) = \int_{y_0}^y \mu(z)dz$, for some $y_0 \in \mathbb{R}$. Then by Itô's formula, $\mathbb{P}_x^\epsilon$ almost surely,

$$\int_0^1 \mu(X_u)dX_u = G(X_1) - G(x) - \frac{\epsilon}{2}\int_0^1 \mu'(X_u)du. \tag{7.19}$$

Consequently ζ_1 becomes

$$\zeta_1 = \exp\left(G(X_1) - G(x) - \frac{\epsilon}{2}\int_0^1 \left\{\mu'(X_u) + \mu(X_u)^2\right\} du\right). \tag{7.20}$$

Overall, equation (7.18) becomes

$$\hat{\mathbb{Q}}_x^{y,\epsilon}[A] = \frac{p_\epsilon(1,x,y)}{q_\epsilon(1,x,y)}\, e^{G(y)-G(x)}\, \hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-\frac{\epsilon}{2}\int_0^1\{\mu'(X_u)+\mu(X_u)^2\}du}\mathbb{1}_{\{X\in A\}}\right]. \tag{7.21}$$

For $\epsilon > 0$, let $\Phi_\epsilon^{(1)}$ and $\widetilde{\Phi}_\epsilon^{(1)}$ be defined by

$$\frac{p_\epsilon(1,x,y)e^{G(y)-G(x)}}{q_\epsilon(1,x,y)} = 1 + \epsilon\Phi_\epsilon^{(1)}(x,y) \tag{7.22}$$

and

$$\hat{\mathbb{E}}_x^{y,\epsilon}\left[\exp\left(-\frac{\epsilon}{2}\int_0^1\left\{\mu'(X_u)+\mu(X_u)^2\right\}du\right)\mathbb{1}_{\{X\in A\}}\right] = \hat{\mathbb{P}}_x^{y,\epsilon}[A]\Big(1+\epsilon\widetilde{\Phi}_\epsilon^{(1)}(x,y;A)\Big), \tag{7.23}$$

respectively. Combining (7.21), (7.22) and (7.23), one obtains

$$\hat{\mathbb{Q}}_x^{y,\epsilon}[A] = \hat{\mathbb{P}}_x^{y,\epsilon}[A](1+\epsilon\Phi_\epsilon^{(1)}(x,y))(1+\epsilon\widetilde{\Phi}_\epsilon^{(1)}(x,y;A)). \tag{7.24}$$

Finally, for an absolutely continuous function $\gamma : [0,1] \longrightarrow \mathbb{R}$ with $\gamma(0) = x$ and $\gamma(1) = y$, let $J = J_{x,y}$ denote the rate function defined by

$$J(\gamma) = \frac{1}{2}\int_0^1 \dot{\gamma}_u^2 du. \tag{7.25}$$

The following result holds.

Proposition 7.2.1.2. *Assume that μ is differentiable and μ' is locally Lipschitz continuous on $\mathbb{R}$. Then the following two properties hold.*

(i) If $\gamma_\cdot^{(x,y)}$ denotes the path joining x and y travelled at constant speed, then

$$\lim_{\epsilon\to 0}\Phi_\epsilon^{(1)}(x,y) = \frac{1}{2}\int_s^1(\mu'+\mu^2)(\gamma_u^{(x,y)})du. \tag{7.26}$$

(ii) Let $A \in \mathcal{F}_1$ be a set of paths. Assume that there exists a unique path $\rho_\cdot = \rho_\cdot^{(x,y,A)}$ such that

$$J(\rho) = \inf_{\phi\in A^\circ} J(\phi) = \inf_{\phi\in\bar{A}} J(\phi), \tag{7.27}$$

where J is defined as in (7.25). Then

$$\lim_{\epsilon\to 0}\widetilde{\Phi}_\epsilon^{(1)}(x,y;A) = -\frac{1}{2}\int_0^1(\mu'+\mu^2)(\rho_u^{(x,y,A)})du. \tag{7.28}$$

The limits appearing above are uniform for (x,y) in a compact subset of $\mathbb{R}^2$.

Proof. A concise proof is given in [6]. See the proof of Lemma 4.3 therein. □

If $A = A_h = \{\sup_{0\le u\le t} X_u \ge h\}$, the previous result can be considered as a first order approximation of the quantity

$$\mathbb{P}_x[\tau_h \le \epsilon \,|\, X_\epsilon = y], \tag{7.29}$$

with respect to ϵ. As we mentioned in the introduction to this chapter, the preceding proposition can be generalized by partial differential equation techniques in order to obtain higher order terms in the expansion of hitting time probabilities for pinned diffusions. This will be done in the sequel. Our analysis will culminate in Section 7.5, where the overall result will be stated. But before we embark on the analysis of higher order terms, let us calculate the first order terms explicitly.

7.2.2 Concrete asymptotics of the first hitting times of pinned diffusions

One Barrier

By means of Proposition 7.2.1.2 we are able to state a result about the asymptotic behavior of first hitting times of pinned diffusions. The path $\gamma_\cdot^{(x,y)}$ mentioned in Proposition 7.2.1.2 (i) consists of a line segment joining x and y, i.e.

$$\gamma_u^{(x,y)} = x + u(y-x), \quad u \in [0,1]. \tag{7.30}$$

Let $A_h(x,y) \in \mathcal{F}_1$ be the set

$$A_h(x,y) = \left\{ \gamma \in \mathcal{C}([0,1],\, \mathbb{R}) \middle| \gamma(0) = x,\, \gamma(1) = y,\, \sup_{0 \le s \le 1} \gamma(s) \ge h \right\} \tag{7.31}$$

and let $\rho_\cdot^{(x,h,y)} : [0,1] \to \mathbb{R}$ be the mapping

$$u \longmapsto \begin{cases} x + \frac{u}{t_h}(h-x), & \text{if } u \in [0, t_h), \\ h + \frac{u-t_h}{1-t_h}(y-h), & \text{if } u \in [t_h, 1], \end{cases} \tag{7.32}$$

where $t_h = \frac{h-x}{2h-x-y}$. In other words the function $\rho_\cdot^{(x,h,y)}$ joins x to h by a line segment travelled at constant speed during the time interval $[0, t_h]$ and then it joins h to y linearly during the remaining time interval $[t_h, 1]$. It is not difficult to show that the path $\rho_\cdot^{(x,h,y)}$ minimizes the functional J on the set $A_h(x,y)$. A variational approach is necessary to prove this fact. But basically, the calculations are simple. The proof is omitted here.

With this at hand, we are now able to state the first important result. Note that the proposition below is a special case of Theorem 2.1 in the article of Baldi and Caramellino [6]. In the upcoming sections of the present chapter, we are then going to generalize this result.

Proposition 7.2.2.1. *We consider the process X defined by* (7.3)*, where $\sigma \equiv 1$ and where μ is assumed to satisfy the assumptions of Proposition 7.2.1.2. We set $\beta = \mu^2 + \mu'$. And for $x, h, y \in \mathbb{R}$, $x, y < h$, we define the expressions*

$$\Phi^{(1)}(x,y) = \frac{1}{2} \int_0^1 \beta(\gamma_u^{(x,y)}) du \tag{7.33}$$

and

$$\widetilde{\Phi}^{(1)}(x,h,y) = -\frac{1}{2}\int_0^1 \beta\Big(\rho_u^{(x,h,y)}\Big)du, \tag{7.34}$$

where $\gamma_\cdot^{(x,y)}$ *and* $\rho_\cdot^{(x,h,y)}$ *are defined by* (7.30) *and* (7.32)*, respectively. Then the stopping time* $\tau_h = \inf\{t > 0 \,|\, X_t \geq h\}$ *satisfies*

$$\begin{aligned}&\mathbb{P}_x[\tau_h \leq \epsilon | X_\epsilon = y] \\ &= \exp\left(-\frac{2(h-y)(h-x)}{\epsilon}\right)\Big(1+\epsilon\Phi^{(1)}(x,y)+\epsilon\widetilde{\Phi}^{(1)}(x,h,y)+\epsilon\mathcal{R}_\epsilon(x,h,y)\Big),\end{aligned} \tag{7.35}$$

where the remainder term $\mathcal{R}_\epsilon(x,h,y)$ *converges uniformly to* 0 *as* $\epsilon \to 0$ *on compact subsets of* $\{(x,h,y)\in\mathbb{R}^3 \,|\, x,y \leq h\}$*.*

Proof. Let $\Phi_\epsilon^{(1)}$ and $\widetilde{\Phi}_\epsilon^{(1)}$ be defined by (7.22) and (7.23), respectively. By formula (7.24) we find

$$\begin{aligned}&\mathbb{P}_x[\tau_h \leq \epsilon | X_\epsilon = y] \\ &= \hat{\mathbb{P}}_x^{y,\epsilon}[X \in A_h]\Big(1+\epsilon\Phi_\epsilon^{(1)}(x,y)\Big)\Big(1+\epsilon\widetilde{\Phi}_\epsilon^{(1)}(x,h,y)\Big) \\ &= \exp\left(-\frac{2(h-y)(h-x)}{\epsilon}\right)\Big(1+\epsilon\Phi_\epsilon^{(1)}(x,y)\Big)\Big(1+\epsilon\widetilde{\Phi}_\epsilon^{(1)}(x,h,y)\Big),\end{aligned} \tag{7.36}$$

where the set A_h is defined by (7.31). Note that for a standard Brownian bridge, connecting x to y, the probability of hitting h equals $\exp(-2(h-x)(h-y))$. We find that

$$\hat{\mathbb{P}}_x^{y,\epsilon}[X \in A_h] = \exp\left(-\frac{2(h-y)(h-x)}{\epsilon}\right). \tag{7.37}$$

It can be verified by straightforward calculations that the rate function J takes its unique minimum on A_h in $\rho_\cdot^{(x,h,y)}$. Now, the result follows directly from Proposition 7.2.1.2. □

We end this paragraph with a remark.

Remark 7.2.2.2. Let $\beta = \mu' + \mu^2$ be integrable. The integral of β along $\gamma_\cdot^{(x,y)}$ is

$$\begin{aligned}\int_0^1 \beta(\gamma_u^{(x,y)})du &= \int_0^1 \beta(x+u(y-x))du \\ &= \frac{1}{y-x}\int_x^y \beta(v)dv.\end{aligned} \tag{7.38}$$

Moreover, one obtains

$$\int_0^1 \beta(\rho_u^{(x,h,y)})du = \int_0^{t_h} \beta\left(x+\frac{u}{t_h}(h-x)\right)du + \int_{t_h}^1 \beta\left(h+\frac{u-t_h}{1-t_h}(y-h)\right)du$$

$$= \frac{t_h}{h-x}\int_x^h \beta(u)du + \frac{1-t_h}{y-h}\int_h^y \beta(u)du$$
$$= \frac{\int_x^h \beta(u)du + \int_y^h \beta(u)du}{2h-x-y}. \tag{7.39}$$

Two Barriers

For the sake of completeness let us briefly describe the asymptotics for two barrier hitting times of pinned diffusions. In the sequel, we will not be concerned with this case any more. Subsequent sections will be dedicated to the analysis of the one barrier problem. We already introduced the path $\gamma_\cdot^{(x,y)}$ in (7.30). Here, we set

$$A_{[l,h]}(x,y) = \left\{\gamma \in \mathcal{C}^0([0,1]) \middle| \gamma(0)=x,\ \gamma(1)=y,\ \sup_{0\le s\le 1}\gamma(s) \ge h,\ \inf_{0\le s\le 1}\gamma(s)\le l\right\}, \tag{7.40}$$

and we define another path $\rho_\cdot^{(x,h,l,y)} : [0,1] \to \mathbb{R}$ that consists of three line segments. The first joins x to h during the interval

$$[0,t_h] = \left[0, \frac{h-x}{2h-2l-x+y}\right], \tag{7.41}$$

the second joins h to l during the time interval

$$[t_h,t_l] = \left[\frac{h-\xi}{2h-2l-x+y}, \frac{2h-l-x}{2h-2l-x+y}\right], \tag{7.42}$$

and the third joins l to y during $[t_l,1]$. In other words, $\rho_\cdot^{(x,h,l,y)}$ is given by

$$u \longmapsto \begin{cases} x + \frac{u}{t_h}(h-x), & \text{if } u \in [0,t_h), \\ h + \frac{u-t_h}{t_l-t_h}(l-h), & \text{if } u \in [t_h,t_l], \\ l + \frac{u-t_l}{1-t_l}(y-l), & \text{if } u \in [t_l,1]. \end{cases} \tag{7.43}$$

It follows by straightforward calculations that $\rho_\cdot^{(x,h,l,y)}$ minimizes the rate function J, which is defined by (7.25), on the set $A_{[l,h]}(x,y)$. But J has another minimum on the set $A_{[l,h]}(x,y)$. For symmetric reasons, the second one is attained in the path $\overleftarrow{\rho}_\cdot^{(x,h,l,y)}$: $[0,1] \to \mathbb{R}$ that first joins x to l by a line segment during the interval

$$[0,\bar{t}_l] = \left[0, \frac{-l+x}{2h-2l-x+y}\right], \tag{7.44}$$

then joins h to l linearly during the time interval

$$[\bar{t}_l,\bar{t}_h] = \left[\frac{-l+x}{2h-2l-x+y}, \frac{-2l+h+x}{2h-2l-x+y}\right], \tag{7.45}$$

and finally joins l to y during $[\bar{t}_h, 1]$ by a third line segment. In a nutshell, the path $\overleftarrow{\rho}_{\cdot}^{(x,h,l,y)}$ is defined by

$$u \longmapsto \begin{cases} x + \frac{u}{\bar{t}_l}(l-x), & \text{if } u \in [0, \bar{t}_l), \\ l + \frac{u-\bar{t}_l}{\bar{t}_h - \bar{t}_l} h - l), & \text{if } u \in [\bar{t}_l, \bar{t}_h], \\ h + \frac{u-\bar{t}_h}{1-\bar{t}_h}(y-h), & \text{if } u \in [\bar{t}_h, 1]. \end{cases} \tag{7.46}$$

It can be verified by direct calculations that, on the one hand,

$$\int_0^1 \beta(\rho_u^{(x,h,l,y)})du = \int_0^1 \beta(\overleftarrow{\rho}_u^{(x,h,l,y)})du, \tag{7.47}$$

and that, on the other hand,

$$J\left(\rho_{\cdot}^{(x,h,l,y)}\right) = J\left(\overleftarrow{\rho}_{\cdot}^{(x,h,l,y)}\right). \tag{7.48}$$

The proof is simple. It is omitted here.

Before we are able to state the main result of this paragraph, some additional considerations are necessary. Let $x, h, l, y \in \mathbb{R}$ with $l < x, y < h$. We want to find an expression for the probability of a Brownian bridge, that connects x to y and that is rescaled by $\sqrt{\epsilon}$, to stay within the interval (l, h). Let $l < a \leq b < h$. Then

$$\int_a^b p_{\sigma^2}^{(l,h)}(t,x,z)dz = \frac{1}{h-l}\int_a^b p_1^{(0,1)}\left(\frac{\sigma^2 t}{(h-l)^2}, \frac{x-l}{h-l}, \frac{z-l}{h-l}\right)dz, \tag{7.49}$$

where

$$p_1^{(0,1)}(t,x,z) = \sum_{k=1}^{\infty} 2\exp(-k^2\pi^2 t/2)\sin(k\pi x)\sin(k\pi z). \tag{7.50}$$

This expression and its relevance were already discussed in Paragraph 3.4.1 of Chapter 3. Since

$$\int_a^b \sin\left(k\pi\frac{z-l}{h-l}\right)dz = \frac{h-l}{k\pi}\left\{\cos\left(k\pi\frac{a-l}{h-l}\right) - \cos\left(k\pi\frac{b-l}{h-l}\right)\right\}, \tag{7.51}$$

one obtains

$$\begin{aligned} &\mathbb{P}_x[l < \sigma L_t^B,\ \sigma H_t^B < h,\ a \leq \sigma B_t \leq b] \\ &= \sum_{k=1}^{\infty} \frac{2}{k\pi}\exp\left(-\frac{k^2\pi^2\sigma^2 t}{2(h-l)^2}\right)\left\{\cos\left(k\pi\frac{a-l}{h-l}\right) - \cos\left(k\pi\frac{b-l}{h-l}\right)\right\}\sin\left(k\pi\frac{x-l}{h-l}\right), \end{aligned} \tag{7.52}$$

where $L_t^B = \inf_{0\le s\le t} B_s$ and $H_t^B = \sup_{0\le s\le t} B_s$. Now, for Brownian motion $(\sigma B_t,\, t\ge 0)$,

$$\mathbb{P}_x[a \le \sigma B_t \le b] = \frac{1}{2}\left(\mathrm{Erf}\left[\frac{b-x}{\sqrt{2\sigma^2 t}}\right] - \mathrm{Erf}\left[\frac{a-x}{\sqrt{2\sigma^2 t}}\right]\right). \tag{7.53}$$

The error function $\mathrm{Erf}(\cdot)$ was introduced in formula (5.51) in Section 5.2.2. Moreover, let us note that

$$2\,\frac{(h-l)}{k\pi}\frac{\cos\left(k\pi\frac{a-l}{h-l}\right)-\cos\left(k\pi\frac{b-l}{h-l}\right)}{\mathrm{Erf}\left[\frac{b-x}{\sqrt{2\sigma^2t}}\right]-\mathrm{Erf}\left[\frac{a-x}{\sqrt{2\sigma^2t}}\right]} \longrightarrow \exp\left(\frac{(a-x)^2}{2\sigma^2 t}\right)\sqrt{2\pi\sigma^2 t}\sin\left(k\pi\frac{a-l}{h-l}\right), \tag{7.54}$$

as $b\to a$. Overall, we obtain the probability of a Brownian bridge, connecting x to y, to remain in the interval (l,h). It is given by the following expression

$$\begin{aligned}
&\mathbb{P}_x[l<\sigma L_t^B,\ \sigma H_t^B<h \,|\, B_t=y]\\
&=\lim_{\delta\to 0}\frac{\mathbb{P}_x[l<\sigma L_t^B,\ \sigma H_t^B<h,\ y\le\sigma B_t\le y+\delta]}{\mathbb{P}_x[a\le\sigma B_t\le b]}\\
&=\sum_{k=1}^{\infty}\frac{2\sqrt{2\pi\sigma^2 t}}{(h-l)}\exp\left(-\frac{k^2\pi^2\sigma^2 t}{2(h-l)^2}+\frac{(y-x)^2}{2\sigma^2 t}\right)\sin\left(k\pi\frac{x-l}{h-l}\right)\sin\left(k\pi\frac{y-l}{h-l}\right).
\end{aligned} \tag{7.55}$$

We made use of the following fact: if $U_\delta(y)$ denotes a ball of radius δ centered around y, then the measure $\mathbb{P}_x[\,\cdot\,|\,B_1\in U_\delta(y)]$ converges weakly to the law of a Brownian bridge from x to y as $\delta\to 0$. For a proof, see e.g. Billingsley [15], p. 101 ff. By formula (7.55), we obtain for $\epsilon>0$,

$$\begin{aligned}
&\hat{\mathbb{P}}_x^{y,\epsilon}\left[l<\inf_{0\le t\le 1}X_t,\ \sup_{0\le t\le 1}X_t<h\right]\\
&=\sum_{k=1}^{\infty}\frac{2\sqrt{2\pi\epsilon}}{(h-l)}\exp\left(-\frac{k^2\pi^2\epsilon}{2(h-l)^2}+\frac{(y-x)^2}{2\epsilon}\right)\sin\left(k\pi\frac{x-l}{h-l}\right)\sin\left(k\pi\frac{y-l}{h-l}\right).
\end{aligned} \tag{7.56}$$

The latter equation follows directly from the fact that $\hat{\mathbb{P}}_x^{y,\epsilon}$ is the law of the process W_x^ϵ pinned at $W_x^\epsilon(1)=y$. Recall the definition of W_x^ϵ in (7.11).

We are now able to state the next result. It describes a special case of Theorem 2.2 in the paper of Baldi and Caramellino [6].

Proposition 7.2.2.3. *Again, we consider a process X defined by* (7.3)*, where $\sigma\equiv 1$ and where μ is assumed to satisfy the assumptions of Proposition 7.2.1.2. And again, let $\beta=\mu^2+\mu'$. For $x,h,l,y\in\mathbb{R}$, $l\le x,y\le h$, we define the expressions*

$$\Phi^{(1)}(x,y)=\frac{1}{2}\int_0^1\beta(\gamma_u^{(x,y)})du \tag{7.57}$$

and

$$\widetilde{\Phi}^{(1)}(x,h,l,y) = -\frac{1}{2}\int_0^1 \beta\Big(\rho_u^{(x,h,l,y)}\Big)du, \tag{7.58}$$

where $\gamma_\cdot^{(x,y)}$ *and* $\rho_\cdot^{(x,h,l,y)}$ *are defined by* (7.30) *and* (7.43)*, respectively. Then the first exit time* $\tau_{[l,h]} = \inf\{t > 0 \,|\, X_t \notin (l,h)\}$ *satisfies*

$$\begin{aligned}&\mathbb{P}_x[\tau_{[l,h]} \le \epsilon | X_\epsilon = y] \\ &= \mathcal{P}(\epsilon,x,h,l,y) \times \Big(1 + \epsilon\Phi^{(1)}(x,y) + \epsilon\widetilde{\Phi}^{(1)}(x,h,l,y) + \epsilon\mathcal{R}_\epsilon(x,h,l,y)\Big),\end{aligned} \tag{7.59}$$

where

$$\begin{aligned}\mathcal{P}(\epsilon,x,h,l,y) = \exp\left(-\frac{2(h-y)(h-x)}{\epsilon}\right) + \exp\left(-\frac{2(y-l)(x-l)}{\epsilon}\right) \\ - \hat{\mathbb{P}}_x^{y,\epsilon}\left[l < \inf_{0\le t\le 1} X_t,\ \sup_{0\le t\le 1} X_t < h\right].\end{aligned} \tag{7.60}$$

The term $\hat{\mathbb{P}}_x^{y,\epsilon}\Big[l < \inf_{0\le t\le 1} X_t,\ \sup_{0\le t\le 1} X_t < h\Big]$ *is given by* (7.56) *and, eventually, the remainder term* $\mathcal{R}_\epsilon(x,h,l,y)$ *converges uniformly to* 0 *as* $\epsilon \to 0$ *on compact subsets of* $\{(x,h,l,y) \in \mathbb{R}^4 \,|\, l \le x, y \le h\}$*.*

Proof. Again, let $\Phi_\epsilon^{(1)}$ and $\widetilde{\Phi}_\epsilon^{(1)}$ be defined by (7.22) and (7.23), respectively. This time, by formula (7.24) and by formula (7.56), we find

$$\begin{aligned}&\mathbb{P}_x[\tau_{[l,h]} \le \epsilon | X_\epsilon = y] \\ &= \hat{\mathbb{P}}_x^{y,\epsilon}[X \in A_{[l,h]}(x,y)]\Big(1 + \epsilon\Phi_\epsilon^{(1)}(x,y)\Big)\Big(1 + \epsilon\widetilde{\Phi}_\epsilon^{(1)}(x,h,l,y)\Big),\end{aligned} \tag{7.61}$$

where the set $A_{[l,h]}$ is defined by (7.40). Clearly $\hat{\mathbb{P}}_x^{y,\epsilon}[X \in A_{[l,h]}(x,y)]$ coincides with the probability that the process W_x^ϵ, pinned at $W_x^\epsilon(1) = y$, crosses both boundaries of (l,h) and, consequently, it equals (7.60). It can be verified by straightforward calculations that, on the set $A_{[l,h]}(x,y)$, the rate function J takes two minima. As we have mentioned above, the minimum is attained in the two paths $\rho_\cdot^{(x,h,l,y)}$ and $\overleftarrow{\rho}_\cdot^{(x,h,l,y)}$, defined by (7.43) and (7.46). Because of (7.47) and (7.48) the presence of two minimum arguments is no restriction. The result now follows directly from Proposition 7.2.1.2. □

We close this section with a trivial observation, which is stated in the next remark.

Remark 7.2.2.4. The integral of the function β along the path $\gamma_\cdot^{(x,y)}$ has already been calculated, see formula (7.38). Moreover, we obtain the following result

$$\begin{aligned}\int_0^1 \beta(\rho_u^{(x,h,l,y)})du = \int_0^{t_h} \beta\left(x + \frac{u}{t_h}(h-x)\right)du + \int_{t_h}^{t_l} \beta\left(h + \frac{u-t_h}{t_l-t_h}(l-h)\right)du \\ + \int_{t_l}^1 \beta\left(l + \frac{u-t_l}{1-t_l}(y-l)\right)du\end{aligned}$$

$$= \frac{t_h}{h-x}\int_x^h \beta(u)du + \frac{t_l - t_h}{h-l}\int_h^l \beta(u)du + \frac{1-t_l}{y-l}\int_l^y \beta(u)du$$

$$= \frac{\int_x^h \beta(u)du + \int_l^h \beta(u)du + \int_l^y \beta(u)du}{2h-2l-x+y}. \tag{7.62}$$

It is straightforward to show that this expression coincides with $\int_0^1 \beta(\overleftarrow{\rho}_u^{(x,h,l,y)})du$.

7.3 Functionals of Brownian bridges

Let $t > 0$ and let $\Omega = \mathcal{C}([0,t],\mathbb{R})$ denote the space of continuous paths and define the coordinate process on Ω by $X_s(w) = w_s$. From X one obtains the filtration $(\mathcal{F}_s,\, 0 \le s \le t)$. For $0 \le s \le t$, the σ-algebra $\mathcal{F}_s$ is defined by $\mathcal{F}_s = \sigma(X_u,\, 0 \le u \le s)$. Let $x, y \in \mathbb{R}$ and let us consider the following bridge process

$$x + \frac{u}{t}(y-x) + \sqrt{\epsilon}\left(B_u - \frac{u}{t}B_t\right), \quad u \in [0,t], \tag{7.63}$$

where B denotes the standard Brownian motion of $\mathbb{R}$. The law of the process (7.63) on $(\Omega, \mathcal{F}_t)$ is denoted with $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ and the corresponding expectation operator is denoted with $\hat{\mathbb{E}}_x^{y,\epsilon,t}$. We will sometimes write $\hat{\mathbb{P}}_x^{y,\epsilon}$ instead of $\hat{\mathbb{P}}_x^{y,\epsilon,1}$ and $\hat{\mathbb{E}}_x^{y,\epsilon}$ instead of $\hat{\mathbb{E}}_x^{y,\epsilon,1}$.

Consider a sufficiently smooth function $\beta : \mathbb{R} \to \mathbb{R}$ and let $h \in \mathbb{R}$ with $x, y \le h$. Baldi and Caramellino [6] made use of large deviation techniques to show that the function $\tilde{\phi}_1^\epsilon$ defined by

$$\hat{\mathbb{E}}_x^{y,\epsilon}\left[\exp\left(-\frac{\epsilon}{2}\int_0^1 \beta(X_u)du\right)\mathbb{1}_{\left\{\sup_{0\le u\le 1} X_u \ge h\right\}}\right] = 1 + \epsilon\tilde{\phi}_1^\epsilon(1,x,h,y), \tag{7.64}$$

satisfies

$$\begin{aligned}&\lim_{\epsilon\to 0}\tilde{\phi}_1^\epsilon(t,x,h,y)\\ &= \int_0^{\frac{h-x}{2h-x-y}} \beta(x+u(2h-x-y))du + \int_{\frac{h-x}{2h-x-y}}^1 \beta(y+(1-u)(2h-x-y))du,\end{aligned} \tag{7.65}$$

where convergence is uniform for (x,h,y) on compact subsets of $\mathbb{R}^3$. This result can be considered as a first order expansion of the left hand side of (7.64) with respect to ϵ. The shortcoming of this method is that it does not provide information about higher order terms, although slight modifications of Baldi and Caramellino's estimates show that the subsequent term in the expansion must belong to $O(\epsilon^2)$. Note that this is not a priori clear, since the supremum of a standard Brownian bridge rescaled by $\sqrt{\epsilon}$ is involved.

The aim is to find higher order expansions of (7.64). The reason why we are interested in this term, is that (7.64) is the key component of the following expression

$$\mathbb{P}_x[\tau_h \leq \epsilon \,|\, X_\epsilon = y], \tag{7.66}$$

where $\tau_h = \inf\{t > 0 \,|\, X_t \geq h\}$. Compare formula (7.18) in the previous paragraph. In a nutshell: if we are able to find an expansion of (7.64), we are able to calculate an expansion of $\mathbb{P}_x[\tau_h \leq \epsilon \,|\, X_\epsilon = y]$. Furthermore, an expansion of this hitting time probability for pinned diffusions will then enable us to calculate an expansion (with respect to $\sqrt{\epsilon}$) of the quantity

$$\mathbb{P}_x \left[\sup_{0 \leq s \leq \epsilon} X_s < h,\ X_\epsilon \in A \right], \tag{7.67}$$

for a diffusion process X and for a Borel-set $A \in \mathcal{B}(\mathbb{R})$ that satisfies $\sup A \leq h$. This is immediately clear from the formula

$$\mathbb{P}_x \left[\sup_{0 \leq s \leq \epsilon} X_s < h,\ X_\epsilon \in A \right] = \int_A \mathbb{P}_x \left[\tau_h > \epsilon \,|\, X_\epsilon = y\right] \mathbb{P}_x[X_\epsilon \in dy] \tag{7.68}$$

and from the fact that there are various expansions of the transition probability density p of X. We have already mentioned that possible references for expansions of p are Aït-Sahalia, [4] and [3], or Kampen [42]. Despite this very concrete justification, finding an expansion of the functional (7.64) with respect to ϵ is an interesting problem in itself.

Structure of the upcoming paragraphs Let us briefly outline our proceeding in the present and in the subsequent sections. In the present Section 7.3, we derive the necessary foundations to determine an overall expansion of the left hand side of (7.64) with respect to ϵ. First, in Paragraph 7.3.1, we assume that there exists a solution to a particular partial differential equation and we show that this solution coincides with the expression

$$\hat{\mathbb{E}}_x^{y,\epsilon,t} \left[\exp\left(-\frac{\epsilon}{2} \int_0^t \beta(X_u) du \right) \right]. \tag{7.69}$$

Starting from this result, we use a PDE approach to derive a system of functions $\phi_k(t, x, y)$, $k \in \mathbb{N}$. These functions can be calculated recursively and the way they are defined suggests that they are candidates – as it were – for the coefficients in the expansion (with respect to ϵ) of the quantity (7.69). For more details see particularly Theorem 7.3.1.7.

In the next step, in Paragraph 7.3.2, a similar procedure shows that

$$\hat{\mathbb{E}}_x^{y,\epsilon,t} \left[\exp\left(-\frac{\epsilon}{2} \int_0^t \beta(X_u) du \right) \mathbb{1}_{\left\{ \sup_{0 \leq u \leq 1} X_u < h \right\}} \right] \tag{7.70}$$

is characterized by a partial differential equation with boundary condition. An approach, analogous to the one of Paragraph 7.3.1, results in another system of recursively defined functions $\tilde{\phi}_k(t,x,h,y)$, $k \in \mathbb{N}$. The new functions $\tilde{\phi}_k(t,x,h,y)$ are candidates for the coefficients in the expansion of (7.70). More details about the $\tilde{\phi}_k$ can be found in Theorem 7.3.2.7.

We emphasize that the results of Paragraph 7.3.1 and 7.3.2 rely entirely on the assumption that there exist solutions to the partial differential equation that describe (7.69) and (7.70). Loosely speaking, the purpose of these two paragraphs is to determine the structure of the ϕ_k and the $\tilde{\phi}_k$.

In Section 7.4, we will then show the following fact. On the Assumptions 7.4.1.1 and 7.4.2.1, which imply that the series

$$v(t,x,y) = 1 + \sum_{i=1}^{\infty} \epsilon^k \phi_k(t,x,y) \tag{7.71}$$

and

$$v_h(t,x,y) = 1 + \sum_{i=1}^{\infty} \epsilon^k \tilde{\phi}_k(t,x,h,y), \tag{7.72}$$

and their derivatives with respect to (t,x), converge uniformly, the expressions (7.71) and (7.72) are solutions to the partial differential equations that characterize the quantities (7.69) and (7.70). Consequently, the series $v(t,x,y)$ and $v_h(t,x,h,y)$ coincide with the respective expressions (7.69) and (7.70). We will also present a result which shows that Assumptions 7.4.1.1 and 7.4.2.1 are reasonable. This will justify our methods and our assumptions in hindsight. For more details see Paragraph 7.4.3.

In Section 7.5 we will combine the findings. A suitable transformation of the functions (7.71) and (7.72), evaluated at $t=1$, will yield an expansion of the quantity (7.66).

Finally, let us mention that our procedure is inspired by the techniques of Fleming and James [25]. They found expansions of different Feynman-Kac formulae for general diffusion processes. Therefore, this chapter can also be considered as a generalization of the results in [25] to a situation where the ordinary diffusion process is replaced by a Brownian bridge.

7.3.1 The Brownian bridge I

Let $\epsilon > 0$ and, in order to simplify our analysis, let $y \in \mathbb{R}$ be fixed from now on. We consider the following stochastic differential equation

$$dX_s = \frac{y - X_s}{t-s} ds + \sqrt{\epsilon} dB_s, \quad 0 \le s < t, \; X_0 = x, \tag{7.73}$$

where B denotes the standard Brownian motion of $\mathbb{R}$. An explicit, and hence strong, solution to (7.73) is given by

$$X_u = x + \frac{u}{t}(y-x) + \sqrt{\epsilon}(t-u)\int_0^u \frac{dB_s}{t-s}, \quad u \in [0,t]. \tag{7.74}$$

Note that the previous solution depends on the starting point x as well as on the end point y, which is attained at time t. We briefly sum up some facts about the process X. First, the process $\left(M_u = \int_0^u \frac{dB_s}{t-s},\ u \leq t\right)$ is a continuous martingale with respect to the Brownian filtration and has quadratic variation

$$\langle M\rangle_u = \int_0^u \frac{ds}{(t-s)^2} = \frac{1}{t-u} - \frac{1}{t}, \quad u \in [0,t]. \tag{7.75}$$

Moreover,

$$(t-u)\int_0^u \frac{dB_s}{t-s} \longrightarrow 0, \tag{7.76}$$

almost surely as $u \to t$, and it is straightforward to show that the process

$$(t-u)\int_0^u \frac{dB_s}{t-s}, \quad 0 \leq u \leq t, \tag{7.77}$$

is equal in law to a standard Brownian bridge on the interval $[0,t]$. In other words, the law of the process (7.74) and the law of the following Brownian bridge

$$x + \frac{u}{t}(y-x) + \sqrt{\epsilon}\left(B_u - \frac{u}{t}B_t\right), \quad u \in [0,t], \tag{7.78}$$

connecting x and y during the time interval $[0,t]$, are the same. By the representation (7.78) of the Brownian bridge, it is obvious that the quadratic variation of X given by (7.74) must equal $\langle X\rangle_u = \epsilon u$. And finally, for the sake of completeness, let us mention that the Brownian bridge has the strong Markov property, see e.g. [24]. Thus the operators

$$T_u f(x) = \hat{\mathbb{E}}_x^{y,\epsilon,t}[f(X_u)], \quad u \in [0,t], \tag{7.79}$$

constitute a Feller-semigroup. We will not make use of this fact later on.

According to our previous notations we denote with $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ the law of (7.74) on the canonical path space $\Omega = \mathcal{C}([0,t],\mathbb{R})$. From now on, X will denote the coordinate variable process given by $X_t(\omega) = \omega_t$, for $\omega \in \Omega$. Our aim is to combine different techniques in such a way that one obtains an expansion of the following functional of a

Brownian bridge

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\mathbb{1}_{\{H_t\geq h\}}\right], \tag{7.80}$$

where we set $H_t = \sup_{0\leq u\leq t} X_u$ as usual. But since

$$\begin{aligned}
&\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\right] \\
&= \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\mathbb{1}_{\{H_t\geq h\}}\right] + \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\mathbb{1}_{\{H_t< h\}}\right],
\end{aligned} \tag{7.81}$$

this is tantamount to finding expansions for both

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\right] \tag{7.82}$$

and

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\mathbb{1}_{\{H_t< h\}}\right]. \tag{7.83}$$

In the present section we first concentrate on the problem of determining an expansion for (7.82). Let us state an interesting result.

Proposition 7.3.1.1. *Fix $y \in \mathbb{R}$ and let $t \in \mathbb{R}_+$. On $\Omega = \mathcal{C}([0,t],\mathbb{R})$ we consider the law of the Brownian Bridge connecting x to y during $[0,t]$ which is defined by (7.74). We denote this law with $\hat{\mathbb{P}}_x^{y,\epsilon,t}$. The corresponding expectation operator is denoted with $\hat{\mathbb{E}}_x^{y,\epsilon,t}$. Consider a Lipschitz continuous and bounded function $\beta : \mathbb{R} \to \mathbb{R}$ and let $\phi_1^\epsilon(t,x)$ be defined by the equation*

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\right] = 1 + \epsilon\phi_1^\epsilon(t,x). \tag{7.84}$$

Moreover, let $\gamma_\cdot = \gamma_\cdot^{(t,x,y)} : [0,t] \to \mathbb{R}$ denote the path $\gamma_u = x + \frac{u}{t}(y-x)$, then

$$\lim_{\epsilon\to 0}\phi_1^\epsilon(t,x) = -\frac{1}{2}\int_0^t \beta(\gamma_u)du. \tag{7.85}$$

Convergence is uniform for (t,x) on compact subsets of $\mathbb{R}_+ \times \mathbb{R}$.

Proof. Let us write

$$\mathcal{D}_\epsilon = \frac{1}{\epsilon}\left(e^{-\frac{\epsilon}{2}\int_0^t \beta(X_u)du} - 1 + \frac{\epsilon}{2}\int_0^t \beta(X_u)du\right) + \frac{1}{2}\int_0^t [\beta(X_u) - \beta(\gamma_u)]du. \tag{7.86}$$

Since β is bounded on $\mathbb{R}$, we have the estimate

$$|\mathcal{D}_\epsilon| \leq \frac{1}{\epsilon}\frac{\epsilon^2}{4}t^2K^2e^{\frac{\epsilon}{2}K} + \frac{1}{2}\int_0^t |\beta(X_u) - \beta(\gamma_u)|du, \tag{7.87}$$

with a suitable constant $K > 0$. Thus,

$$\limsup_{\epsilon \to 0} \hat{\mathbb{E}}_x^{y,\epsilon,t}[|\mathcal{D}_\epsilon|] \leq \frac{1}{2} \limsup_{\epsilon \to 0} \hat{\mathbb{E}}_x^{y,\epsilon,t} \int_0^t |\beta(X_u) - \beta(\gamma_u)| du. \tag{7.88}$$

Now fix $\delta > 0$ and let us denote with $U_\delta(\gamma)$ the open ball on $\mathcal{C}([0,t],\mathbb{R})$, being centered around the path γ and having radius δ. By the Lipschitz property and the boundedness of β, for a suitable constant K, the following estimate holds

$$\hat{\mathbb{E}}_x^{y,\epsilon,t} \int_0^t |\beta(X_u) - \beta(\gamma_u)| du \leq K\delta + 2K \hat{\mathbb{P}}_x^{y,\epsilon,t}[U_\delta(\gamma)^c]. \tag{7.89}$$

Let $\bar{\gamma}_\delta$ be the path minimizing the rate function $J(z) = \frac{1}{2}\int_0^t \dot{z}^2(s)ds$ on the interior and the closure of $U_\delta(\gamma)^c$. Then, by a standard large deviations argument,

$$\lim_{\epsilon \to 0} \epsilon \log \hat{\mathbb{P}}_x^{y,\epsilon,t}[U_\delta(\gamma)^c] = -J(\bar{\gamma}_\delta), \tag{7.90}$$

which implies

$$\limsup_{\epsilon \to 0} \hat{\mathbb{E}}_x^{y,\epsilon,t} \int_0^t |\beta(X_u) - \beta(\gamma_u)| du \leq K\delta. \tag{7.91}$$

The constant δ was chosen arbitrarily and thus the statement follows. Finally, note that by our estimates the limits appearing above are uniform for (t,x) on compact subsets of $\mathbb{R}_+ \times \mathbb{R}$. $\square$

The latter proposition makes use of large deviation principles. It is not satisfactory inasmuch as it does not provide information about higher order terms. The next theorem, which is inspired by the techniques of Fleming and James [25], renders higher order expansions possible. We introduce a notation. Let $m,n \in \mathbb{N}$. Henceforth, we will denote with $\mathcal{C}^{m,n}([0,T] \times \mathbb{R}, \mathbb{R})$ the space of functions $g : [0,T] \times \mathbb{R} \to \mathbb{R}$ that satisfy $g \in \mathcal{C}^{m,n}((0,T) \times \mathbb{R}, \mathbb{R})$ and $g_{k,l} \in \mathcal{C}([0,T] \times \mathbb{R}, \mathbb{R})$, for all partial derivatives $g_{k,l}$ with $k \leq m$ and $l \leq n$.

Theorem 7.3.1.2. *Fix $y \in \mathbb{R}$ and $T \in \mathbb{R}_+$. Let $t \in [0,T]$. Again, on $\Omega = \mathcal{C}([0,t], \mathbb{R})$, let $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ denote the law of the Brownian bridge connecting x to y during $[0,t]$, which is defined by (7.74), and let $\hat{\mathbb{E}}_x^{y,\epsilon,t}$ denote the corresponding expectation operator. Let $\beta \in \mathcal{C}_b(\mathbb{R},\mathbb{R})$ and let $v \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}, \mathbb{R})$. If v is a solution to the partial differential equation*

$$\begin{aligned}&\frac{\partial}{\partial t} v(t,x) \\ &= \frac{\epsilon}{2}\frac{\partial^2}{\partial x^2} v(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x} v(t,x) - \frac{\epsilon}{2}\beta(x) v(t,x), \quad \forall t \in (0,T],\ \forall x \in \mathbb{R},\end{aligned} \tag{7.92}$$

with initial condition

$$v(0,x) = 1, \quad \forall x \in \mathbb{R}, \tag{7.93}$$

then v is given by

$$v(t,x) = \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds\right)\right], \quad (t,x) \in [0,T] \times \mathbb{R}. \tag{7.94}$$

Remark 7.3.1.3. In the above theorem, we assumed that $\beta : \mathbb{R} \to \mathbb{R}$ is a bounded function. This assumption is not necessarily needed, but it helps us to avoid technical difficulties. In order to prove Theorem 7.3.1.2, we have to make use of Itô's formula. The term $\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds\right)$ appears and the boundedness assumption guarantees that the involved expectations exist. Later on, when deriving overall expansions, we will be able to weaken the assumption of boundedness. The assumption $v \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}, \mathbb{R})$ also seems to be very restrictive. This assumption guarantees that the partial derivatives $(t,x) \mapsto v_{0,i}(t,x)$, $i = 1,2$, do not explode for $t = 0$. The upcoming results, especially those of Section 7.4, will show that, indeed, there are no singularities with respect to the time variable t. This will justify our assumption in hindsight.

Proof (of Theorem 7.3.1.2). In order to show (7.94), we apply Itô's formula to the function

$$u \longmapsto \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u, X_u), \quad u \in [0,t]. \tag{7.95}$$

The result is

$$\begin{aligned} d&\left\{\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u,X_u)\right\} \\ &\qquad = -\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u,X_u)\frac{\epsilon}{2}\beta(X_u)du \\ &\qquad - \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{1,0}(t-u,X_u)du \\ &\qquad + \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u,X_u)dX_u \\ &\qquad + \frac{1}{2}\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,2}(t-u,X_u)d\langle X\rangle_u. \end{aligned} \tag{7.96}$$

The dynamics of X are given by the stochastic differential equation (7.73). Also recall that $d\langle X\rangle_u = \epsilon\, du$. Additionally, v was assumed to solve (7.92). Hence the previous equation becomes

$$\begin{aligned} d&\left\{\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u,X_u)\right\} \\ &\qquad = \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u,X_u)\sqrt{\epsilon}dB_u. \end{aligned} \tag{7.97}$$

Integrating from 0 to t with respect to u and taking expectations, one obtains the assertion. □

We are now able to derive a first order expansion of the right hand side of (7.94) with respect to ϵ. The result is stated in the next theorem.

Theorem 7.3.1.4. *We fix $y \in \mathbb{R}$ and $T \in \mathbb{R}_+$. Let $\beta \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ and let $v \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}, \mathbb{R})$ be a solution to (7.92) with initial condition (7.93). Define $\phi_1^\epsilon : [0,T] \times \mathbb{R} \to \mathbb{R}$ by the equation*

$$v(t,x) = 1 + \epsilon \phi_1^\epsilon(t,x), \quad \forall t \in [0,T], \forall x \in \mathbb{R}. \tag{7.98}$$

Then, for all $(t,x) \in (0,T] \times \mathbb{R}$, the function ϕ_1^ϵ solves the differential equation

$$\frac{\partial}{\partial t}\phi_1^\epsilon(t,x) = \frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}\phi_1^\epsilon(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\phi_1^\epsilon(t,x) - \frac{\epsilon}{2}\beta(x)\phi_1^\epsilon(t,x) - \frac{1}{2}\beta(x), \tag{7.99}$$

with initial condition

$$\phi_1^\epsilon(0,x) = 0, \quad \forall x \in \mathbb{R}. \tag{7.100}$$

Moreover, ϕ_1^ϵ has the following representation

$$\phi_1^\epsilon(t,x) = -\frac{1}{2}\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\beta(X_u)du\right]. \tag{7.101}$$

Here, $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ denotes the law of (7.74) and $\hat{\mathbb{E}}_x^{y,\epsilon,t}$ denotes the corresponding expectation operator. Finally, $\lim_{\epsilon\to 0}\phi_1^\epsilon(t,x) = \phi_1(t,x)$, uniformly for (t,x) on compact subsets of $[0,T] \times \mathbb{R}$, where the function ϕ_1 is defined by

$$\phi_1(t,x) = -\frac{1}{2}\int_0^t \beta(\gamma_u)du = -\frac{1}{2}\int_0^t \beta\left(x + \frac{u}{t}(y-x)\right)du. \tag{7.102}$$

Proof. By definition we have

$$\phi_1^\epsilon(t,x) = \frac{v(t,x)-1}{\epsilon}. \tag{7.103}$$

Straightforward calculations show that, if v is a solution to (7.92), the function ϕ_1^ϵ must satisfy the differential equation (7.99) with the initial condition

$$\phi_1^\epsilon(0,x) = 0. \tag{7.104}$$

We apply Itô's formula to the function

$$u \longmapsto \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\phi_1^\epsilon(t-u, X_u) + \int_0^u \exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds, \tag{7.105}$$

in order to obtain

$$
\begin{aligned}
& d\left\{\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\phi_1^\epsilon(t-u,X_u)+\int_0^u \exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds\right\} \\
= & \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_a)da\right)\beta(X_u)du \\
- & \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\phi_1^\epsilon(t-u,X_u)\frac{\epsilon}{2}\beta(X_u)du \\
- & \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\phi_{1;\,1,0}^\epsilon(t-u,X_u)du \\
+ & \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\phi_{1;\,0,1}^\epsilon(t-u,X_u)dX_u \\
+ & \frac{\epsilon}{2}\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\phi_{1;\,0,2}^\epsilon(t-u,X_u)du.
\end{aligned}
\tag{7.106}
$$

The process X satisfies the stochastic differential equation (7.73). Hence, by integrating from 0 to t and by taking expectations, we see that an explicit solution to the partial differential equation (7.99) satisfies the Feynman-Kac formula for the Brownian bridge given by (7.101). Recall that β was assumed to be bounded. Thus, by dominated convergence,

$$\lim_{\epsilon\to 0}\phi_1^\epsilon(t,x)=\phi_1(t,x)=-\frac{1}{2}\int_0^t \beta\left(x+\frac{s}{t}(y-x)\right)ds=-\frac{1}{2}\int_0^t \beta\left(\gamma_s\right)ds. \tag{7.107}$$

Obviously, convergence is uniform for (t,x) on compact subsets of $[0,T]\times\mathbb{R}$. This concludes the proof of the theorem. □

By letting ϵ tend to 0 in (7.99) one might be tempted to guess that ϕ_1 satisfies the differential equation

$$\frac{\partial}{\partial t}\phi_1(t,x)=\frac{y-x}{t}\frac{\partial}{\partial x}\phi_1(t,x)-\frac{1}{2}\beta(x). \tag{7.108}$$

Indeed, from the next proposition, which holds in a quite general context, this property can be derived.

Proposition 7.3.1.5. *Let $g:[0,T]\times\mathbb{R}\to\mathbb{R}$ be an integrable function. Moreover, let us assume that g is continuously differentiable with respect to both variables and let $\gamma_\cdot=\gamma_\cdot^{(t,x,y)}:[0,t]\to\mathbb{R}$ denote the path $\gamma_u=x+\frac{u}{t}(y-x)$. The function*

$$\phi(t,x)=\int_0^t g(t-u,\gamma_u)du \tag{7.109}$$

satisfies, for all $(t,x)\in(0,T]\times\mathbb{R}$, the differential equation

$$\frac{\partial}{\partial t}\phi(t,x)=\frac{y-x}{t}\frac{\partial}{\partial x}\phi(t,x)+g(t,x). \tag{7.110}$$

Remark 7.3.1.6. Of course, the latter proposition remains true, if the function g does not depend on the time parameter t.

Proof. First, let us note that

$$\frac{\partial}{\partial t}\int_0^t g(t-u,\gamma_u)du = g(0,y) - \frac{(y-x)}{t^2}\int_0^t u\, g_{0,1}\left(t-u,\gamma_u\right)du$$
$$+\int_0^t g_{1,0}\left(t-u,\gamma_u\right)\, du \tag{7.111}$$

and

$$\frac{y-x}{t}\frac{\partial}{\partial x}\int_0^t g(t-u,\gamma_u)du = \frac{(y-x)}{t^2}\int_0^t (t-u)g_{0,1}(t-u,\gamma_u)\, du. \tag{7.112}$$

Therefore,

$$\frac{\partial}{\partial t}\int_0^t g(t-u,\gamma_u)du - \frac{y-x}{t}\frac{\partial}{\partial x}\int_0^t g(t-u,\gamma_u)du$$
$$= g(0,y) + \int_0^t g_{1,0}\left(t-u,\gamma_u\right)\, du - \frac{(y-x)}{t}\int_0^t g_{0,1}\left(t-u,\gamma_u\right)du. \tag{7.113}$$

Since

$$\frac{\partial}{\partial u} g\left(t-u,\gamma_u\right) = -g_{1,0}\left(t-u,\gamma_u\right)\, du + \frac{(y-x)}{t} g_{0,1}\left(t-u,\gamma_u\right), \tag{7.114}$$

one obtains

$$\frac{\partial}{\partial t}\left\{\int_0^t g(t-u,\gamma_u)du\right\} - \frac{y-x}{t}\frac{\partial}{\partial x}\left\{\int_0^t g(t-u,\gamma_u)du\right\} = g(t,x), \tag{7.115}$$

which completes the proof of the proposition. □

By an approach similar to the one of Theorem 7.3.1.4 we obtain a higher order expansion with respect to ϵ. The result is given in the next theorem.

Theorem 7.3.1.7. *Again, we fix $y \in \mathbb{R}$ and $T \in \mathbb{R}_+$. Let $n \in \mathbb{N}$ and assume that $\beta \in \mathcal{C}_b^{2n}(\mathbb{R},\mathbb{R})$. Moreover, we assume that β and its derivatives are uniformly bounded on $\mathbb{R}$. We denote with $v \in \mathcal{C}^{1,2n}([0,T]\times\mathbb{R},\mathbb{R})$ a solution to the partial differential equation* (7.92) *with initial condition* (7.93). *As before, let $\gamma_u = \gamma_u^{(t,x,y)} = x + \frac{u}{t}(y-x)$, $u \in [0,t]$. For $(t,x) \in [0,T]\times\mathbb{R}$, let ϕ_n^ϵ be defined by*

$$v(t,x) = 1 + \epsilon\phi_1(t,x) + \ldots + \epsilon^{n-1}\phi_{n-1}(t,x) + \epsilon^n\phi_n^\epsilon(t,x), \tag{7.116}$$

where we set $\phi_0 \equiv 1$ and, for $1 \le k < n$, the functions ϕ_k are supposed to satisfy the recursion scheme

$$\phi_k(t,x) = \int_0^t g_{k-1}(t-u,\gamma_u)du, \tag{7.117}$$

with

$$g_{k-1}(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{k-1}(t,x) - \frac{1}{2}\beta(x)\phi_{k-1}(t,x). \tag{7.118}$$

Then, for all $(t,x) \in (0,T] \times \mathbb{R}$, *the function* ϕ_n^ϵ *satisfies the differential equation*

$$\frac{\partial}{\partial t}\phi_n^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{n-1}^\epsilon(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\phi_n^\epsilon(t,x) - \frac{1}{2}\beta(x)\phi_{n-1}^\epsilon(t,x), \tag{7.119}$$

with initial condition

$$\phi_n^\epsilon(0,x) = 0, \quad \forall x \in \mathbb{R}. \tag{7.120}$$

It has the alternative representation

$$\phi_n^\epsilon(t,x) = \int_0^t g_{n-1}^\epsilon(t-u,\gamma_u)du, \tag{7.121}$$

where g_{n-1}^ϵ *denotes the function*

$$g_{n-1}^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{n-1}^\epsilon(t,x) - \frac{1}{2}\beta(x)\phi_{n-1}^\epsilon(t,x). \tag{7.122}$$

Finally, $\lim_{\epsilon\to 0}\phi_n^\epsilon(t,x) = \phi_n(t,x)$, *uniformly for* (t,x) *on compact subsets of* $[0,T]\times\mathbb{R}$.

Proof. By definition, we have

$$\phi_1^\epsilon(t,x) = \frac{v(t,x)-1}{\epsilon}. \tag{7.123}$$

And iteratively we have, for $n \geq 1$,

$$\phi_n^\epsilon(t,x) = \frac{\phi_{n-1}^\epsilon(t,x) - \phi_{n-1}(t,x)}{\epsilon}. \tag{7.124}$$

In order to demonstrate the ideas behind the proof, let us consider the case $n = 2$. This means, we first focus on the term ϕ_2^ϵ. Recall that ϕ_1 satisfies

$$\frac{\partial}{\partial t}\phi_1(t,x) = \frac{y-x}{t}\frac{\partial}{\partial x}\phi_1(t,x) - \frac{1}{2}\beta(x), \tag{7.125}$$

with initial condition

$$\phi_1(0,x) = 0. \tag{7.126}$$

A proof for this fact was given in Proposition 7.3.1.5. Thus, for $n = 2$, formula (7.124) directly implies that $\phi_2^\epsilon(t,x)$ must solve

$$\frac{\partial}{\partial t}\phi_2^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_1^\epsilon(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\phi_2^\epsilon(t,x) - \frac{1}{2}\beta(x)\phi_1^\epsilon(t,x), \tag{7.127}$$

with initial condition

$$\phi_2^\epsilon(0,x) = 0. \tag{7.128}$$

Conversely, let us assume that ϕ_2^ϵ denotes a solution to the differential equation (7.127). If one applies Itô's formula to the function

$$u \longmapsto \phi_2^\epsilon(t-u,\gamma_u) + \int_0^u g_1^\epsilon(t-s,\gamma_s)ds, \quad u \in [0,t], \tag{7.129}$$

where g_1^ϵ is defined by

$$g_1^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_1^\epsilon(t,x) - \frac{1}{2}\beta(x)\phi_1^\epsilon(t,x), \tag{7.130}$$

then one obtains

$$\begin{aligned} & d\left\{\phi_2^\epsilon(t-u,\gamma_u) + \int_0^u g_1^\epsilon(t-s,\gamma_s)ds\right\} \\ & = -\phi_{2;\,1,0}^\epsilon(t-u,\gamma_u) + \phi_{2;\,0,1}^\epsilon(t-u,\gamma_u)\frac{y-\gamma_u}{t-u}du + g_1^\epsilon(t-u,\gamma_u). \end{aligned} \tag{7.131}$$

Note, that in the latter equation we made use of the fact that

$$\frac{\partial \gamma_u}{\partial u} = \frac{y-x}{t} = \frac{y-\gamma_u}{t-u}. \tag{7.132}$$

Integrating (7.131) from 0 to t, we see that a solution to (7.127) is given by

$$\phi_2^\epsilon(t,x) = \int_0^t g_1^\epsilon(t-u,\gamma_u)du. \tag{7.133}$$

The explicit formula (7.101) enables us to calculate directly the second derivative of

$$\phi_1^\epsilon(t,x) = -\frac{1}{2}\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\beta(X_u)du\right] \tag{7.134}$$

with respect to x. Note that our assumptions about β allow for an interchange of integration and differentiation. This can be directly deduced from Satz 2, Chapter 11 in the book of Forster [27]. Then, by the boundedness of β and its derivatives in combination with dominated convergence, one can easily infer that

$$\lim_{\epsilon\to 0} g_1^\epsilon(t,x) = g_1(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_1(t,x) - \frac{1}{2}\beta(x)\phi_1(t,x). \tag{7.135}$$

Convergence in the previous formula is clearly uniform for (t,x) on compact subsets of $[0,T]\times\mathbb{R}$. And thus, by (7.133) in combination with the dominated convergence theorem, we find that

$$\lim_{\epsilon \to 0} \phi_2^\epsilon(t,x) = \int_0^t g_1(t-u,\gamma_u)du, \tag{7.136}$$

uniformly for (t,x) on compact subsets of $[0,T] \times \mathbb{R}$.

The proof for $n > 2$ works in an analogous way. We only give a sketch. Note that, for $k < n$, the functions ϕ_k, defined by the recursion (7.117), (7.118), satisfy the differential equation

$$\frac{\partial}{\partial t}\phi_k(t,x) = \frac{y-x}{t}\frac{\partial}{\partial x}\phi_k(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{k-1}(t,x) - \frac{1}{2}\beta(x)\phi_{k-1}(t,x), \tag{7.137}$$

with initial condition

$$\phi_k(0,x) = 0. \tag{7.138}$$

By the definition of ϕ_n^ϵ, see (7.124), and by the fact that ϕ_2^ϵ satisfies the differential equation (7.127), one iteratively obtains that the function ϕ_n^ϵ must satisfy

$$\frac{\partial}{\partial t}\phi_n^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{n-1}^\epsilon(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\phi_n^\epsilon(t,x) - \frac{1}{2}\beta(x)\phi_{n-1}^\epsilon(t,x), \tag{7.139}$$

with initial condition

$$\phi_n^\epsilon(0,x) = 0. \tag{7.140}$$

Again, by a Feynman-Kac approach similar to the one above, we infer that a solution to (7.139) is given by (7.121) and (7.122). The function g_{n-1}^ϵ in (7.122) recursively depends on the expression

$$\phi_1^\epsilon(t,x) = -\frac{1}{2}\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\beta(X_u)du\right] \tag{7.141}$$

and its derivatives up to the order $2n$. Due to the explicit representation of ϕ_1^ϵ, we are able to calculate these derivatives directly. Thus, the same arguments that we applied to prove the case $n = 2$ can be used to show that, for $n \geq 2$,

$$\lim_{\epsilon \to 0} \phi_n^\epsilon(t,x) = \phi_n(t,x) = \int_0^t g_{n-1}(t-u,\gamma_u)du, \tag{7.142}$$

uniformly for (t,x) on compact subsets of $[0,T] \times \mathbb{R}$. □

The next corollary states a result that we already mentioned in the previous proof. It is a trivial consequence of Proposition 7.3.1.5. But because of its importance it shall be stressed again.

Corollary 7.3.1.8. *Let $\beta \in \mathcal{C}_b^{2n}(\mathbb{R},\mathbb{R})$ with bounded derivatives of all orders, and let $\phi_0 \equiv 1$. The functions $\phi_k : [0,T] \times \mathbb{R} \to \mathbb{R}$, $k = 1,\ldots,n$, which are recursively defined* *by* (7.117) *and* (7.118)*, are solutions to the following partial differential equation*

$$\frac{\partial}{\partial t}\phi_k(t,x) = \frac{y-x}{t}\frac{\partial}{\partial x}\phi_k(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{k-1}(t,x) - \frac{1}{2}\beta(x)\phi_{k-1}(t,x), \tag{7.143}$$

for all $(t,x) \in (0,T] \times \mathbb{R}$*, with initial condition*

$$\phi_k(0,x) = 0, \quad \forall x \in \mathbb{R}. \tag{7.144}$$

Proof. The result is a direct consequence of Proposition 7.3.1.5. □

In order to assess the results of this paragraph, let us note again that a little calculus shows that the function

$$\phi_1(t,x) = -\frac{1}{2}\int_0^t \beta\left(\gamma_s\right) ds = -\frac{1}{2}\int_0^t \beta\left(x + \frac{s}{t}(y-x)\right) ds \tag{7.145}$$

solves the differential equation

$$\frac{\partial}{\partial t}\phi_1(t,x) = \frac{y-x}{t}\frac{\partial}{\partial x}\phi_1(t,x) - \frac{1}{2}\beta(x)\underbrace{\phi_0(t,x)}_{=1}. \tag{7.146}$$

The function ϕ_1 evidently satisfies the initial condition

$$\phi_1(0,x) = 0, \quad \forall x \in \mathbb{R}. \tag{7.147}$$

Consequently, the first order expansion we found by means of PDE techniques perfectly matches the result derived by means of large deviations theory. Compare the statement of Proposition 7.3.1.1. But additionally, our sophisticated approach enabled us to calculate higher order terms and thus, we found an overall expansion of the quantity

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_s)ds}\right], \tag{7.148}$$

with respect to ϵ.

7.3.2 The Brownian bridge II - boundary conditions

Before we are going to plunge into the details, let us define two functions that play a crucial role within this section. We have already defined the path $\gamma_\cdot = \gamma_\cdot^{(t,x,y)} : [0,t] \to \mathbb{R}$ by

$$\gamma_u = x + \frac{u}{t}(y-x), \quad u \in [0,t]. \tag{7.149}$$

Now, for $h \geq x, y$, define $\rho_{\cdot} = \rho_{\cdot}^{(t,x,h,y)} : [0,t] \to \mathbb{R}$ by

$$\rho_u = \begin{cases} x + \frac{u}{t}(2h - x - y) & , \quad \text{if} \quad 0 \leq u \leq t\frac{h-x}{2h-x-y}, \\ y + \frac{t-u}{t}(2h - x - y) & , \quad \text{if} \quad t\frac{h-x}{2h-x-y} < u \leq t. \end{cases} \tag{7.150}$$

The function $\gamma_{\cdot}$ connects x to y linearly during the interval [0,t], whereas the function $\rho_{\cdot}$ first goes from x to h linearly during the interval

$$\left[0, t\frac{h-x}{2h-x-y}\right], \tag{7.151}$$

and then connects h to y linearly during the interval

$$\left[t\frac{h-x}{2h-x-y}, t\right]. \tag{7.152}$$

Note that the functions $\gamma_{\cdot}$ and $\rho_{\cdot}$ minimize the functional

$$J(z) = \frac{1}{2}\int_0^t \dot{z}^2(s)ds \tag{7.153}$$

within the classes

$$\left\{z \in \mathcal{C}([0,t], \mathbb{R}) \;\middle|\; z(0) = x, \;z(t) = y\right\} \tag{7.154}$$

and

$$\left\{z \in \mathcal{C}([0,t], \mathbb{R}) \;\middle|\; z(0) = x, \;z(t) = y, \;\sup_{0\leq u\leq t} z(u) \geq h\right\}, \tag{7.155}$$

respectively. We state the first result of this section.

Theorem 7.3.2.1. *Let $h, y \in \mathbb{R}$ with $y \leq h$. We assume that both y and h are fixed. Again, on $\Omega = \mathcal{C}([0,t], \mathbb{R})$, let $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ denote the law of the Brownian bridge connecting x to y during $[0,t]$, which is defined by (7.74), and let $\hat{\mathbb{E}}_x^{y,\epsilon,t}$ denote the corresponding expectation operator. If $\beta \in \mathcal{C}_b(\mathbb{R},\mathbb{R})$ and if the function $v_h \in \mathcal{C}^{1,2}([0,T] \times (-\infty, h], \mathbb{R})$ solves, for all t $\in (0,T]$ and for all $x \in (-\infty, h]$, the differential equation*

$$\frac{\partial}{\partial t}v_h(t,x) = \frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}v_h(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}v_h(t,x) - \frac{\epsilon}{2}\beta(x)v_h(t,x), \tag{7.156}$$

with the initial condition

$$v_h(0,x) = 1, \quad \forall x \in \mathbb{R}, \tag{7.157}$$

and with the boundary condition

$$v_h(t,h) = 0, \quad \forall t \in [0,T], \tag{7.158}$$

then v_h has the following representation

$$
\begin{aligned}
v_h(t,x) &= \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_u)du}\mathbb{1}_{\{\sup_{0\le u\le t} X_u<h\}}\right] \\
&= \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[e^{-\frac{\epsilon}{2}\int_0^t \beta(X_u)du}\mathbb{1}_{\{\tau_h>t\}}\right], \qquad (7.159)
\end{aligned}
$$

where $\tau_h = \inf\{u>0 \,|\, X_u \ge h\}$.

Proof. The dynamics of X are determined by the stochastic differential equation

$$
dX_u = \frac{y-X_u}{t-u}du + \sqrt{\epsilon}dB_u, \quad u\in[0,t],\ X_0=x. \qquad (7.160)
$$

Moreover, since $v_h(t,h)=0$, for all $t\in[0,T]$, the following equality holds

$$
e^{-\frac{\epsilon}{2}\int_0^t \beta(X_u)du}\mathbb{1}_{\{\sup_{0\le u\le t} X_u<h\}} = e^{-\frac{\epsilon}{2}\int_0^{t\wedge\tau_h} \beta(X_u)du} v_h(t-t\wedge\tau_h,\, X_{t\wedge\tau_h}). \qquad (7.161)
$$

Clearly, τ_h is a stopping time with respect to $\mathcal{F}_t=\sigma(X_s,\, s\le t)$. Thus, applying Itô's formula to the function

$$
u \longmapsto e^{-\frac{\epsilon}{2}\int_0^u \beta(X_u)du} v_h(t-u, X_u), \quad u\in[0,t], \qquad (7.162)
$$

integrating from 0 to $t\wedge\tau_h$ and then taking expectations, one obtains the result. Further details are omitted here, since the formulae are basically the same as in the proof of Theorem 7.3.1.2, compare especially formula (7.96). □

By the previous theorem, we found a representation for a solution v_h to the differential equation (7.156). Our next aim will be to find an expansion of v_h. A comparison with the results of Baldi and Caramellino - see [6] or the discussion in Section 7.2 - makes the following ansatz seem reasonable

$$
v(t,x)-v_h(t,x) = \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left(1+\sum_{i=1}^{n-1}\epsilon^i\tilde{\phi}_i(t,x)+\epsilon^n\tilde{\phi}_n^\epsilon(t,x)\right). \qquad (7.163)
$$

Here v denotes a solution of (7.92) without boundary condition, this means a solution to the problem we considered in the previous Paragraph 7.3.1. The $\tilde{\phi}_i$ are the functions to be determined and it turns out, that this approach is the right one. From Lemma 4.3 in [6] we know that, for the case $n=1$, we have $\lim_{\epsilon\to 0}\tilde{\phi}_1^\epsilon(t,x)=\tilde{\phi}_1(t,x)$, uniformly for (t,x) on compact subsets of $[0,T]\times(-\infty,h]$, where

$$
\tilde{\phi}_1(t,x) = -\frac{1}{2}\int_0^t \beta\left(\rho_u\right)du. \qquad (7.164)
$$

This result was derived by making use of large deviations techniques. Let us again examine the term $\tilde{\phi}_1^\epsilon$. The next theorem gives a PDE result for $\tilde{\phi}_1^\epsilon$. Higher order terms are treated later on.

Theorem 7.3.2.2. *Let $h, y \in \mathbb{R}$, $y \leq h$, and $T \in \mathbb{R}_+$ be fixed. We assume that $\beta \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ and let $v \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}, \mathbb{R})$ denote a solution of* (7.92) *with initial condition* (7.93)*. Moreover, let $v_h \in \mathcal{C}^{1,2}([0,T] \times (-\infty, h], \mathbb{R})$ be a solution to* (7.156) *with initial condition* (7.157) *and with the boundary condition* (7.158)*. Let $\tilde{\phi}_1^\epsilon$ be defined by the equation*

$$v(t,x) - v_h(t,x) = \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left(1 + \epsilon\tilde{\phi}_1^\epsilon(t,x)\right). \tag{7.165}$$

Then $\tilde{\phi}_1^\epsilon(t,x)$ satisfies, for all $t \in (0,T]$ and for all $x \in (-\infty, h]$, the differential equation

$$\begin{aligned}&\frac{\partial}{\partial t}\tilde{\phi}_1^\epsilon(t,x)\\ &= \frac{1}{2}\epsilon\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) - \frac{1}{2}\epsilon\beta(x)\tilde{\phi}_1^\epsilon(t,x) - \frac{1}{2}\beta(x),\end{aligned} \tag{7.166}$$

with initial condition

$$\tilde{\phi}_1^\epsilon(0,x) = 0, \quad \forall x \in \mathbb{R}, \tag{7.167}$$

and with the boundary condition

$$\tilde{\phi}_1^\epsilon(t,h) = \phi_1^\epsilon(t,h), \quad \forall t \in [0,T]. \tag{7.168}$$

Remark 7.3.2.3. Note that v and v_h satisfy the same differential equation. Compare (7.92) and (7.156). But they differ inasmuch as v_h additionally satisfies the boundary condition (7.158).

Proof (of Theorem 7.3.2.2). Define $V(t,x) = v(t,x) - v_h(t,x)$. By solving (7.165) for $\tilde{\phi}_1^\epsilon(t,x)$, one obtains

$$\tilde{\phi}_1^\epsilon(t,x) = \frac{\exp\left(2\frac{(h-x)(h-y)}{\epsilon t}\right)V(t,x) - 1}{\epsilon}. \tag{7.169}$$

Differentiation with respect to t yields

$$\frac{\partial}{\partial t}\tilde{\phi}_1^\epsilon(t,x) = \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\left(-V(t,x)\frac{2(h-x)(h-y)}{t^2\epsilon} + \frac{\partial}{\partial t}V(t,x)\right). \tag{7.170}$$

And differentiation with respect to x yields

$$\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) = \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\left(-V(t,x)\frac{2(h-y)}{t\epsilon} + \frac{\partial}{\partial x}V(t,x)\right), \tag{7.171}$$

whereas the second derivative with respect to x is given by

$$\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) = \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\left(V(t,x)\frac{4(h-y)^2}{t^2\epsilon^2} - \frac{4(h-y)}{t\epsilon}\frac{\partial}{\partial x}V(t,x) + \frac{\partial^2}{\partial x^2}V(t,x)\right). \tag{7.172}$$

Since both v and v_h satisfy the differential equation (7.92), one obtains

$$\begin{aligned}\frac{\partial}{\partial t}\tilde{\phi}_1^\epsilon(t,x) &= -\frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{2(h-x)(h-y)}{t^2\epsilon}V(t,x)\\ &+ \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\left\{\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}V(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}V(t,x) - \frac{\epsilon}{2}\beta(x)V(t,x)\right\}.\end{aligned} \tag{7.173}$$

Let us solve (7.170), (7.171) and (7.172) for $\frac{\partial}{\partial t}V(t,x)$, $\frac{\partial}{\partial x}V(t,x)$ and $\frac{\partial^2}{\partial x^2}V(t,x)$, respectively, and let us plug the resulting expressions into (7.173). The calculations are routine and yield

$$\begin{aligned}\frac{\partial}{\partial t}\tilde{\phi}_1^\epsilon(t,x) &= -\frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{2(h-x)(h-y)}{t^2\epsilon}V(t,x)\\ &+ \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\left\{\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}V(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}V(t,x) - \frac{\epsilon}{2}\beta(x)V(t,x)\right\}\\ &= -\frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{2(h-x)(h-y)}{t^2\epsilon}V(t,x)\\ &+ \frac{1}{2}\epsilon\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) - e^{\frac{2(h-x)(h-y)}{t\epsilon}}\left(V(t,x)\frac{2(h-y)^2}{t^2\epsilon^2} - \frac{2(h-y)}{t\epsilon}\frac{\partial}{\partial x}V(t,x)\right)\\ &+ \frac{y-x}{t}\left(\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) + \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}V(t,x)\frac{2(h-y)}{t\epsilon}\right)\\ &- \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{\epsilon}{2}\beta(x)V(t,x)\\ &= -\frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{2(h-x)(h-y)}{t^2\epsilon}V(t,x)\\ &+ \frac{1}{2}\epsilon\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) - e^{\frac{2(h-x)(h-y)}{t\epsilon}}\left(V(t,x)\frac{2(h-y)^2}{t^2\epsilon^2}\right)\\ &+ \frac{2(h-y)}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) + \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{4(h-y)^2}{t^2\epsilon}V(t,x)\\ &+ \frac{y-x}{t}\left(\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) + \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}V(t,x)\frac{2(h-y)}{t\epsilon}\right)\\ &- \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{\epsilon}{2}\beta(x)V(t,x)\\ &= \frac{1}{2}\epsilon\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) + \frac{2(h-y)}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\end{aligned}$$

$$
\begin{aligned}
&- \frac{e^{\frac{2(h-x)(h-y)}{t\epsilon}}}{\epsilon}\frac{\epsilon}{2}\beta(x)V(t,x)\\
&= \frac{1}{2}\epsilon\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) + \frac{2(h-y)}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\\
&- \frac{1}{2}\epsilon\beta(x)\tilde{\phi}_1^\epsilon(t,x) - \frac{1}{2}\beta(x).
\end{aligned}
\tag{7.174}
$$

This computation shows that $\tilde{\phi}_1^\epsilon$ satisfies the differential equation (7.166). The initial condition and the boundary condition are obvious from the definitions of v and ϕ_1^ϵ, respectively. □

Remark 7.3.2.4. We formally let $\epsilon \to 0$ in the differential equation (7.166). A heuristical argument suggests that $\tilde{\phi}_1(t,x)$ must satisfy

$$
\frac{\partial}{\partial t}\tilde{\phi}_1(t,x) = \frac{2(h-y)}{t}\frac{\partial}{\partial x}\tilde{\phi}_1(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}\tilde{\phi}_1(t,x) - \frac{1}{2}\beta(x), \tag{7.175}
$$

with the initial condition

$$
\tilde{\phi}_1(0,x) = 0, \tag{7.176}
$$

and with the boundary condition

$$
\tilde{\phi}_1(t,h) = \phi_1(t,h). \tag{7.177}
$$

The next theorem and the subsequent proposition show that this is correct, indeed.

Theorem 7.3.2.5. *Let $y,h \in \mathbb{R}$, $y < h$, and $T \in \mathbb{R}_+$ be fixed and let us assume that $\beta \in \mathcal{C}_b(\mathbb{R},\mathbb{R})$. Moreover, let $\tilde{\phi}_1^\epsilon \in \mathcal{C}^{1,2}([0,T]\times(-\infty,h],\mathbb{R})$ denote a solution of the partial differential equation* (7.166) *with initial condition* (7.167) *and boundary condition* (7.168)*. We consider the following Brownian bridge, that connects x to $2h-y$ during the interval $[0,t]$,*

$$
x + \frac{u}{t}(2h-x-y) + \sqrt{\epsilon}\left(B_u - \frac{u}{t}B_t\right), \quad u \in [0,t]. \tag{7.178}
$$

Here, B denotes the standard Brownian motion of $\mathbb{R}$. On $\Omega = \mathcal{C}([0,t],\mathbb{R})$ we consider the law of (7.178)*, which we denote with $\hat{\mathbb{P}}_x^{2h-y,\epsilon,t}$. If τ_h denotes the first hitting time of the coordinate variable process X with h, i.e. $\tau_h = \inf\{u \ge 0 \,|\, X_u \ge h\}$, then $\tilde{\phi}_1^\epsilon$ has the representation*

$$
\begin{aligned}
\tilde{\phi}_1^\epsilon(t,x) = &-\frac{1}{2}\hat{\mathbb{E}}_x^{2h-y,\epsilon,t}\left[\int_0^{\tau_h}\exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds\right]\\
&+ \hat{\mathbb{E}}_x^{2h-y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^{\tau_h}\beta(X_s)ds\right)\phi_1^\epsilon(t-\tau_h,h)\right],
\end{aligned}
\tag{7.179}
$$

where ϕ_1^ϵ is the function (7.101)*. Finally, $\tilde{\phi}_1^\epsilon(t,x) \to \tilde{\phi}_1(t,x)$, uniformly for (t,x) on* *compact subsets of* $[0,T]\times(-\infty,h]$ *as* $\epsilon\to 0$.

Proof. The dynamics of (7.178) are described by the stochastic differential equation

$$dX_u = \frac{2h - X_u - y}{t-u}du + \sqrt{\epsilon}dB_u, \quad u\in[0,t],\ X_0 = x, \tag{7.180}$$

where B denotes the standard Brownian motion of $\mathbb{R}$. Applying Itô's formula to the expression

$$\begin{aligned} u \longmapsto & \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_1^\epsilon(t-u,X_u) \\ & -\frac{1}{2}\int_0^u \exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds, \quad u\in[0,t], \end{aligned} \tag{7.181}$$

one obtains

$$\begin{aligned} & d\left\{\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_1^\epsilon(t-u,X_u) - \frac{1}{2}\int_0^u \exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds\right\} \\ &= -\frac{1}{2}\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_a)da\right)\beta(X_u)du \\ &- \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_1^\epsilon(t-u,X_u)\frac{\epsilon}{2}\beta(X_u)du \\ &- \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_{1;\,1,0}^\epsilon(t-u,X_u)du \\ &+ \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_{1;\,0,1}^\epsilon(t-u,X_u)\frac{2h-X_u-y}{t-u}du \\ &+ \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_{1;\,0,1}^\epsilon(t-u,X_u)\sqrt{\epsilon}dB_u \\ &+ \frac{\epsilon}{2}\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\tilde{\phi}_{1;\,0,2}^\epsilon(t-u,X_u)du. \end{aligned} \tag{7.182}$$

The function $\tilde{\phi}_1^\epsilon$ satisfies the differential equation (7.166). Therefore, by integrating from 0 to τ_h and by taking expectations on both sides of the latter equation, one obtains

$$\begin{aligned} \tilde{\phi}_1^\epsilon(t,x) = &-\frac{1}{2}\hat{\mathbb{E}}_x^{2h-y,\epsilon,t}\left[\int_0^{\tau_h}\exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds\right] \\ &+\hat{\mathbb{E}}_x^{2h-y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^{\tau_h}\beta(X_s)ds\right)\tilde{\phi}_1^\epsilon\left(t-\tau_h,X_{\tau_h}\right)\right]. \end{aligned} \tag{7.183}$$

Under $\hat{\mathbb{P}}_x^{2h-y,\epsilon,t}$ we have $X_{\tau_h}\equiv h$. Consequently, $\tilde{\phi}_1^\epsilon\left(t-\tau_h,h\right) = \phi_1^\epsilon\left(t-\tau_h,h\right)$ a.s. $\hat{\mathbb{P}}_x^{2h-y,\epsilon,t}$, and we obtain

$$\begin{aligned} \tilde{\phi}_1^\epsilon(t,x) = &-\frac{1}{2}\hat{\mathbb{E}}_x^{2h-y,\epsilon,t}\left[\int_0^{\tau_h}\exp\left(-\frac{\epsilon}{2}\int_0^s \beta(X_a)da\right)\beta(X_s)ds\right] \\ &+\hat{\mathbb{E}}_x^{2h-y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^{\tau_h}\beta(X_s)ds\right)\phi_1^\epsilon\left(t-\tau_h,h\right)\right]. \end{aligned} \tag{7.184}$$

Let $\bar{\rho}^{(t,x,h,y)}_{\cdot} = \bar{\rho}_{\cdot} : [0,t] \to \mathbb{R}$ denote the linear path from x to $2h-y$, given by

$$\bar{\rho}_u = x + \frac{u}{t}(2h-x-y), \quad u \in [0,t]. \tag{7.185}$$

If ϵ tends to 0, X_u tends to $\bar{\rho}_u$ and τ_h tends to $t(h-x)/(2h-x-y)$, a.s. $\hat{\mathbb{P}}^{2h-y,\epsilon,t}_x$. Therefore,

$$\begin{aligned}
\lim_{\epsilon\to 0} \tilde{\phi}^{\epsilon}_1(t,x) &= -\frac{1}{2}\int_0^{t\frac{h-x}{2h-x-y}} \beta(\bar{\rho}_s)ds + \phi_1\left(t - t\frac{h-x}{2h-x-y}, h\right) \\
&= -\frac{1}{2}\int_0^{t\frac{h-x}{2h-x-y}} \beta(\bar{\rho}_s)ds - \frac{1}{2}\int_0^{t\frac{h-y}{2h-x-y}} \beta\left(h + \frac{s}{t}(2h-x-y)\right)ds \\
&= -\frac{1}{2}\int_0^t \beta(\rho_s)ds = \tilde{\phi}_1(t,x),
\end{aligned} \tag{7.186}$$

where convergence is uniform on compact subsets of $[0,T]\times(-\infty,h]$, due to the boundedness and continuity of β. $\square$

We prove a result for the limiting function $\tilde{\phi}_1$ we derived in the previous theorem.

Proposition 7.3.2.6. *Let $\beta \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$. The real valued function*

$$\begin{aligned}
\tilde{\phi}_1(t,x) =& -\frac{1}{2}\int_0^t \beta(\rho_u)du \\
=& -\frac{1}{2}\int_0^{\frac{t(h-x)}{2h-x-y}} \beta\left(x + \frac{u}{t}(2h-x-y)\right)du \\
& -\frac{1}{2}\int_{\frac{t(h-x)}{2h-x-y}}^t \beta\left(y + \frac{t-u}{t}(2h-x-y)\right)du
\end{aligned} \tag{7.187}$$

satisfies, for all $t \in (0,T]$ and for all $x \in (-\infty,h]$, the differential equation (7.175)*. In addition, it satisfies the initial and the boundary conditions*

$$\begin{aligned}
&\tilde{\phi}_1(t,h) = \phi_1(t,h), && \forall t \in [0,T] \\
&\tilde{\phi}_1(0,x) = 0, && \forall x \in (-\infty,h].
\end{aligned} \tag{7.188}$$

Proof. By setting $x = h$, one can directly verify the boundary condition. In this case, $\gamma_{\cdot} \equiv \rho_{\cdot}$. The differential equation can also be verified by direct calculations. We have

$$\begin{aligned}
&\frac{\partial}{\partial t}\int_0^{\frac{t(h-x)}{2h-x-y}} \beta\left(x + \frac{u}{t}(2h-x-y)\right)du \\
&= \frac{(h-x)}{2h-x-y}\beta(h) - \int_0^{t\frac{h-x}{2h-x-y}} \frac{u(2h-x-y)}{t^2}\beta'\left(x + \frac{u(2h-x-y)}{t}\right)du
\end{aligned} \tag{7.189}$$

and

$$\frac{\partial}{\partial x}\int_0^{\frac{t(h-x)}{2h-x-y}}\beta\left(x+\frac{u}{t}(2h-x-y)\right)du$$
$$=-t\frac{(h-y)}{(2h-x-y)^2}\beta(h)+\int_0^{t\frac{(h-x)}{2h-x-y}}\frac{(t-u)\beta'\left(x+\frac{u(2h-x-y)}{t}\right)}{t}du. \tag{7.190}$$

On the other hand, we have

$$\frac{\partial}{\partial t}\int_{\frac{t(h-x)}{2h-x-y}}^{t}\beta\left(y+\frac{t-u}{t}(2h-x-y)\right)du$$
$$=-\frac{(h-x)}{2h-x-y}\beta(h)+\beta(y)+\int_{t\frac{(h-x)}{2h-x-y}}^{t}\frac{u(2h-x-y)\beta'\left(y+\frac{(t-u)(2h-x-y)}{t}\right)}{t^2}du \tag{7.191}$$

and

$$\frac{\partial}{\partial x}\int_{\frac{t(h-x)}{2h-x-y}}^{t}\beta\left(y+\frac{t-u}{t}(2h-x-y)\right)du$$
$$=\frac{t(h-y)}{(2h-x-y)^2}\beta(h)-\int_{\frac{t(h-x)}{2h-x-y}}^{t}\frac{(t-u)\beta'\left(y+\frac{(t-u)(2h-x-y)}{t}\right)}{t}du. \tag{7.192}$$

Finally,

$$\frac{\partial}{\partial u}\beta\left(x+\frac{u}{t}(2h-x-y)\right)=\frac{(2h-x-y)}{t}\beta'\left(x+\frac{u}{t}(2h-x-y)\right) \tag{7.193}$$

and

$$\frac{\partial}{\partial u}\beta\left(y+\frac{t-u}{t}(2h-x-y)\right)=-\frac{(2h-x-y)}{t}\beta'\left(y+\frac{t-u}{t}(2h-x-y)\right). \tag{7.194}$$

Hence,

$$-\frac{1}{2}\frac{\partial}{\partial t}\int_0^t\beta(\rho_u)du+\frac{1}{2}\frac{2h-x-y}{t}\frac{\partial}{\partial x}\int_0^t\beta(\rho_u)du=-\frac{1}{2}\beta(y)+\frac{1}{2}\int_0^t\frac{\partial}{\partial u}\beta(\rho_u)du=-\frac{1}{2}\beta(x), \tag{7.195}$$

which yields the result. □

Now, that we have found a first order expansion, the question is: how can we calculate higher order terms? Theorem 7.3.2.7 below gives the answer to this question.

Theorem 7.3.2.7. *Let $y,h\in\mathbb{R}$, $y<h$, and $T\in\mathbb{R}_+$ be fixed and let $n\in\mathbb{N}$ and $\beta\in\mathcal{C}_b^{2n}(\mathbb{R},\mathbb{R})$. We assume that there exists a function $v\in\mathcal{C}^{1,2n}([0,T]\times\mathbb{R},\mathbb{R})$ that solves* (7.92) *with initial condition* (7.93) *and a function $v_h\in\mathcal{C}^{1,2n}([0,T]\times(-\infty,h),\mathbb{R})$ that*

solves (7.156) *with initial condition* (7.157) *and with boundary condition* (7.158)*. Let the function* $\rho_{\cdot}^{(t,x,h,y)} = \rho_{\cdot}$ *be defined by* (7.150)*. Moreover, set* $\tilde{\phi}_0 \equiv 1$ *and let the functions* $\tilde{\phi}_k$ *be recursively defined via*

$$\tilde{\phi}_k(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{k-1}(t-u,\rho_u)du + \phi_k\left(t\frac{h-y}{2h-x-y}, h\right), \quad \text{for } k \le n, \tag{7.196}$$

where the functions $\tilde{g}_{k-1}$ *are given by*

$$\tilde{g}_{k-1}(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}(t,x), \quad \text{for } k \le n, \tag{7.197}$$

and where the functions ϕ_k *are defined by* (7.117) *and* (7.118)*. Let the function* $\tilde{\phi}_n^\epsilon$ *be defined by the equation*

$$v(t,x) - v_h(t,x) = \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left(1 + \sum_{i=1}^{n-1}\epsilon^i\tilde{\phi}_i(t,x) + \epsilon^n\tilde{\phi}_n^\epsilon(t,x)\right). \tag{7.198}$$

Then $\tilde{\phi}_n^\epsilon$ *satisfies, for all* $(t,x) \in (0,T] \times (-\infty,h]$*, the differential equation*

$$\frac{\partial}{\partial t}\tilde{\phi}_n^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}^\epsilon(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_2^\epsilon(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{n-1}^\epsilon(t,x), \tag{7.199}$$

with initial condition

$$\tilde{\phi}_n^\epsilon(0,x) = 0, \quad \forall x \in (-\infty,h], \tag{7.200}$$

and with boundary condition

$$\tilde{\phi}_n^\epsilon(t,h) = \phi^\epsilon(t,h), \quad \forall t \in [0,T]. \tag{7.201}$$

Moreover, the function $\tilde{\phi}_n^\epsilon$ *has the following representation*

$$\tilde{\phi}_n^\epsilon(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{n-1}^\epsilon(t-u,\rho_u)du + \phi_k^\epsilon\left(t\frac{h-y}{2h-x-y}, h\right), \tag{7.202}$$

where $\tilde{g}_{n-1}^\epsilon$ *denotes the function*

$$\tilde{g}_{n-1}^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}^\epsilon(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{n-1}^\epsilon(t,x). \tag{7.203}$$

If $\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}^\epsilon(t,x)$ *converges uniformly to* $\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}(t,x)$ *on compact subsets of* $[0,T] \times (-\infty,h]$*, then* $\lim_{\epsilon\to 0}\tilde{\phi}_n^\epsilon(t,x) = \tilde{\phi}_n(t,x)$*, uniformly for* (t,x) *on compact subsets of* $[0,T] \times (-\infty,h]$*.*

It is necessary to stress a fact that is outlined in the following remark.

Remark 7.3.2.8. In contrast to the result without boundary condition, which we stated in Theorem 7.3.1.7, we must assume that the function $\tilde{\phi}^{\epsilon}_{n-1}(t,x)$ is twice continuously differentiable with respect to x and that the second derivative $\frac{\partial^2}{\partial x^2}\tilde{\phi}^{\epsilon}_{n-1}(t,x)$ converges uniformly to $\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}(t,x)$ as $\epsilon \to 0$. Similar to the case without boundary conditions, the functions $\tilde{\phi}^{\epsilon}_n$ recursively depend on the function $\tilde{\phi}^{\epsilon}_1$ and its derivatives with respect to x. By the following representation of ϕ^{ϵ}_1,

$$\phi^{\epsilon}_1(t,x) = -\frac{1}{2}\hat{\mathbb{E}}^{y,\epsilon,t}_x\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right)\beta(X_u)du\right], \tag{7.204}$$

it was possible to calculate the derivatives of ϕ^{ϵ}_1 directly and thus, we were able to state uniform convergence. Here, in the case with boundary conditions, it is not possible to differentiate directly, because $\tilde{\phi}^{\epsilon}_1$ depends on the hitting time τ_h in a non-trivial way, see formula (7.179). Consequently, the assumption of uniform convergence of the second derivative $\frac{\partial^2}{\partial x^2}\tilde{\phi}^{\epsilon}_{n-1}$ is necessary. However, in the next section, we will give a criterion that ensures the smoothness of the functions $\tilde{\phi}^{\epsilon}_n$ and implies uniform convergence.

Proof (of Theorem 7.3.2.7). We proceed as we did in the proof of Theorem 7.3.1.7 in the previous section. We iteratively set

$$\tilde{\phi}^{\epsilon}_k(t,x) = \frac{\tilde{\phi}^{\epsilon}_{k-1}(t,x) - \phi_{k-1}(t,x)}{\epsilon}, \quad k \le n. \tag{7.205}$$

For $n=2$, one deduces that the function $\tilde{\phi}^{\epsilon}_2(t,x)$ must satisfy

$$\frac{\partial}{\partial t}\tilde{\phi}^{\epsilon}_2(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}^{\epsilon}_1(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}^{\epsilon}_2(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}^{\epsilon}_1(t,x). \tag{7.206}$$

This follows easily from the definition of $\tilde{\phi}^{\epsilon}_2$ in combination with the fact that $\tilde{\phi}_1$ satisfies the differential equation (7.175). By applying Itô's formula to the function

$$u \longmapsto \tilde{\phi}^{\epsilon}_2(t-u,\rho_u) + \int_0^u g^{\epsilon}_1(t-s,\rho_s)ds, \quad u \in \left[0, t\frac{h-x}{2h-x-y}\right], \tag{7.207}$$

one obtains

$$\begin{aligned} &d\left\{\tilde{\phi}^{\epsilon}_2(t-u,\rho_u) + \int_0^u g^{\epsilon}_1(t-s,\rho_s)ds\right\} \\ &= -\tilde{\phi}^{\epsilon}_{2;\,1,0}(t-u,\rho_u) + \tilde{\phi}^{\epsilon}_{2;\,0,1}(t-u,\rho_u)\frac{2h-\rho_u-y}{t-u}du + \tilde{g}^{\epsilon}_1(t-u,\rho_u). \end{aligned} \tag{7.208}$$

In the previous calculation we made use of the fact that

$$\frac{\partial \rho_u}{\partial u} = \frac{2h-x-y}{t} = \frac{2h-\rho_u-y}{t-u}. \tag{7.209}$$

Integrating (7.208) from 0 to $t(h-x)/(2h-x-y)$, we see that a solution to (7.206) is given by

$$\tilde{\phi}_2^\epsilon(t,x) - \tilde{\phi}_2^\epsilon\left(t - t\frac{h-x}{2h-x-y}, h\right) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_1^\epsilon(t-u,\rho_u)du, \tag{7.210}$$

where g_1^ϵ denotes the function

$$\tilde{g}_1^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_1^\epsilon(t,x). \tag{7.211}$$

Since $\tilde{\phi}_2^\epsilon\left(t - t\frac{h-x}{2h-x-y}, h\right) = \phi_2^\epsilon\left(t - t\frac{h-x}{2h-x-y}, h\right)$, equation (7.210) becomes

$$\begin{aligned}\tilde{\phi}_2^\epsilon(t,x) &= \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_1^\epsilon(t-u,\rho_u)du + \phi_2^\epsilon\left(t - t\frac{h-x}{2h-x-y}, h\right)\\ &= \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_1^\epsilon(t-u,\rho_u)du + \phi_2^\epsilon\left(t\frac{h-y}{2h-x-y}, h\right).\end{aligned} \tag{7.212}$$

Due to our regularity assumptions, we infer that $\lim_{\epsilon\to 0}\tilde{\phi}_2^\epsilon(t,x) = \tilde{\phi}_2(t,x)$, uniformly for (t,x) on compact subsets of $[0,T]\times(-\infty,h]$. By letting ϵ tend to 0 in (7.212) we obtain a function $\tilde{\phi}_2$ that solves the partial differential equation

$$\frac{\partial}{\partial t}\tilde{\phi}_2(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_1(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_2(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_1(t,x). \tag{7.213}$$

Indeed, this is correct. A proof for this fact will be given in Proposition 7.3.2.9 below.

The general case is treated in an analogous way. By arguments similar to the ones above, one iteratively finds that, for $n \geq 2$, the function $\tilde{\phi}_n^\epsilon$ satisfies the following differential equation

$$\frac{\partial}{\partial t}\tilde{\phi}_n^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}^\epsilon(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_n^\epsilon(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{n-1}^\epsilon(t,x), \tag{7.214}$$

and therefore, on the regularity assumptions we made about the functions $\tilde{\phi}_1^\epsilon,\ldots,\tilde{\phi}_{n-1}^\epsilon$, we can derive the representation (7.202). Moreover, $\lim_{\epsilon\to 0}\tilde{\phi}_n^\epsilon(t,x) = \tilde{\phi}_n(t,x)$, uniformly for (t,x) on compact subsets of $[0,T]\times(-\infty,h]$. Finally, as equation (7.214) already suggests, $\tilde{\phi}_n$ satisfies the differential equation

$$\frac{\partial}{\partial t}\tilde{\phi}_n(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_n(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{n-1}(t,x), \tag{7.215}$$

with initial condition

$$\tilde{\phi}_n(0,x) = 0, \quad \forall x \in (-\infty,h], \tag{7.216}$$

and with boundary condition

$$\tilde{\phi}_n(t,h) = \phi_n(t,h), \quad \forall t \in [0,T]. \tag{7.217}$$

Again, this fact follows directly from the statement of Proposition 7.3.2.9 below and consequently, there is nothing left to show. ☐

Let us prove the missing facts in the proof of Theorem 7.3.2.7.

Proposition 7.3.2.9. *For $k \in \mathbb{N}$, let the functions ϕ_k and ϕ_k^ϵ be defined as in Theorem 7.3.1.7 and let the functions $\tilde{\phi}_{k-1}$ and $\tilde{\phi}_{k-1}^\epsilon$ be defined as in Theorem 7.3.2.7. Moreover, let the functions $g_{k-1} : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $g_{k-1}^\epsilon : [0,T] \times \mathbb{R} \to \mathbb{R}$ be defined by*

$$\tilde{g}_{k-1}(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}(t,x), \tag{7.218}$$

and

$$\tilde{g}_{k-1}^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}^\epsilon(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}^\epsilon(t,x), \tag{7.219}$$

respectively. Let $\rho_\cdot^{(t,x,h,y)} = \rho_\cdot$ be the path defined by (7.150). In this situation, the functions

$$\tilde{\phi}_k(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{k-1}(t-u,\rho_u)du + \phi_k\left(t\frac{h-y}{2h-x-y}, h\right) \tag{7.220}$$

and

$$\tilde{\phi}_k^\epsilon(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{k-1}^\epsilon(t-u,\rho_u)du + \phi_k^\epsilon\left(t\frac{h-y}{2h-x-y}, h\right) \tag{7.221}$$

satisfy, for all $(t,x) \in (0,T] \times (-\infty,h]$, the differential equations

$$\frac{\partial}{\partial t}\tilde{\phi}_k^\epsilon(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}^\epsilon(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_k^\epsilon(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}^\epsilon(t,x), \tag{7.222}$$

$$\tilde{\phi}_k^\epsilon(0,x) = 0, \quad \forall x \in (-\infty,h], \tag{7.223}$$

$$\tilde{\phi}_k^\epsilon(t,h) = \phi_k^\epsilon(t,h), \quad \forall t \in [0,T], \tag{7.224}$$

and

$$\frac{\partial}{\partial t}\tilde{\phi}_k(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_k(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}(t,x), \tag{7.225}$$

$$\tilde{\phi}_k(0,x) = 0, \quad \forall x \in (-\infty,h], \tag{7.226}$$

$$\tilde{\phi}_k(t,h) = \phi_k(t,h), \quad \forall t \in [0,T], \tag{7.227}$$

respectively, provided that the involved functions $\tilde{\phi}_{k-1}^\epsilon$ and $\tilde{\phi}_{k-1}$ belong to $\mathcal{C}^{1,2}([0,T] \times \mathbb{R}, \mathbb{R})$.

Proof. By direct calculations, one obtains

$$\begin{aligned}
\frac{\partial}{\partial t}\tilde{\phi}_k^{\epsilon}(t,x) = & \frac{(h-x)}{2h-x-y}\tilde{g}_{k-1}^{\epsilon}\left(t\frac{(h-y)}{2h-x-y},h\right) \\
& - (2h-x-y)\int_0^{\frac{t(h-x)}{2h-x-y}} \frac{u}{t^2}\tilde{g}_{k-1;0,1}^{\epsilon}\left(t-u,x+\frac{u}{t}(2h-x-y)\right)du \\
& + \int_0^{\frac{t(h-x)}{2h-x-y}} \tilde{g}_{k-1;1,0}^{\epsilon}\left(t-u,\frac{u}{t}(2h-x-y)\right)du \\
& + \frac{h-y}{2h-x-y}\phi_{k;\,1,0}^{\epsilon}\left(t\frac{h-y}{2h-x-y},h\right)
\end{aligned} \tag{7.228}$$

and

$$\begin{aligned}
\frac{\partial}{\partial x}\tilde{\phi}_k^{\epsilon}(t,x) = & -\frac{t(h-y)\tilde{g}_{k-1}^{\epsilon}\left(t\frac{(h-y)}{2h-x-y},h\right)}{(2h-x-y)^2} \\
& + \int_0^{\frac{t(h-x)}{2h-x-y}} \frac{(t-u)\tilde{g}_{k-1;0,1}^{\epsilon}\left(t-u,x+\frac{u}{t}(2h-x-y)\right)}{t}du \\
& + t\frac{h-y}{(2h-x-y)^2}\phi_{k;\,1,0}^{\epsilon}\left(t\frac{h-y}{2h-x-y},h\right).
\end{aligned} \tag{7.229}$$

Therefore,

$$\begin{aligned}
& \frac{\partial}{\partial t}\tilde{\phi}_k^{\epsilon}(t,x) - \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_k^{\epsilon}(t,x) \\
& = \tilde{g}_{k-1}^{\epsilon}\left(t\frac{(h-y)}{2h-x-y},h\right) \\
& - \frac{2h-x-y}{t}\int_0^{\frac{t(h-x)}{2h-x-y}} \tilde{g}_{k-1;0,1}^{\epsilon}\left(t-u,x+\frac{u}{t}(2h-x-y)\right)du \\
& + \int_0^{\frac{t(h-x)}{2h-x-y}} \tilde{g}_{k-1;1,0}^{\epsilon}\left(t-u,\frac{u}{t}(2h-x-y)\right)du.
\end{aligned} \tag{7.230}$$

Moreover,

$$\begin{aligned}
& \frac{\partial}{\partial u}\tilde{g}_{k-1}^{\epsilon}\left(t-u,x+\frac{u}{t}(2h-x-y)\right) \\
& = -\tilde{g}_{k-1;1,0}^{\epsilon}\left(t-u,x+\frac{u}{t}(2h-x-y)\right) \\
& + \frac{2h-x-y}{t}\tilde{g}_{k-1;0,1}^{\epsilon}\left(t-u,x+\frac{u}{t}(2h-x-y)\right).
\end{aligned} \tag{7.231}$$

By (7.231), equation (7.230) becomes

$$\frac{\partial}{\partial t}\tilde{\phi}_k^{\epsilon}(t,x) - \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_k^{\epsilon}(t,x)$$

$$= \tilde{g}^{\epsilon}_{k-1}\left(t\frac{(h-y)}{2h-x-y}, h\right) - \int_0^{t\frac{h-x}{2h-x-y}} \frac{\partial}{\partial u}\tilde{g}^{\epsilon}_{k-1}\left(t-u, \rho_u\right) du = \tilde{g}^{\epsilon}_{k-1}(t,x). \tag{7.232}$$

The calculations for $\tilde{\phi}_k$ are the same. Finally, the statements about the boundary conditions are obvious. This shows the result. □

The next corollary is a byproduct of the proof of Theorem 7.3.2.7. It contains statements that have already been proved. But, on account of their importance, they shall be stressed again.

Corollary 7.3.2.10. *Let $n \in \mathbb{N}$ and consider the functions ϕ_n and ϕ^{ϵ}_n defined in the proof of Theorem 7.3.1.7. If $\tilde{\phi}^{\epsilon}_n$ belongs to $\mathcal{C}^{1,2}([0,T]\times(-\infty,h],\mathbb{R})$ and satisfies the differential equation* (7.199) *with initial condition* (7.200) *and boundary condition* (7.201)*, then $\tilde{\phi}^{\epsilon}_n$ has the following representation*

$$\tilde{\phi}^{\epsilon}_n(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}^{\epsilon}_{n-1}(t-u,\rho_u)du + \phi^{\epsilon}_n\left(t\frac{h-y}{2h-x-y}, h\right), \tag{7.233}$$

where the function $\tilde{g}^{\epsilon}_{n-1}$ is given by (7.203)*. Analogously, if the function $\tilde{\phi}_n \in \mathcal{C}^{1,2}([0,T]\times(-\infty,h],\mathbb{R})$ solves the differential equation*

$$\frac{\partial}{\partial t}\tilde{\phi}_n(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{n-1}(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_n(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{n-1}(t,x), \tag{7.234}$$

with initial condition

$$\tilde{\phi}_n(0,x) = 0, \quad \forall x \in (-\infty,h], \tag{7.235}$$

and with boundary condition

$$\tilde{\phi}_n(t,h) = \phi_n(t,h), \quad \forall t \in [0,T], \tag{7.236}$$

then we have

$$\tilde{\phi}_n(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{n-1}(t-u,\rho_u)du + \phi_n\left(t\frac{h-y}{2h-x-y}, h\right), \tag{7.237}$$

where the function $\tilde{g}_{n-1}$ is defined by (7.197)*.*

Proof. A proof of the assertion for $\tilde{\phi}^{\epsilon}_n$ can be found in the proof of Theorem 7.3.2.7. See formula (7.210) for the case $n=2$. The proof for $n \geq 2$ follows in an analogous way. Finally, the calculations for $\tilde{\phi}_n$ are obviously the same. □

We conclude this section with a very important observation.

Remark 7.3.2.11. On the interval $[0, t(h-x)/(2h-x-y)]$ the path $\rho_\cdot = \rho_\cdot^{(t,x,h,y)}$, that consists of the two line segments defined by (7.150), coincides with the linear path

$\bar{\rho}_\cdot = \bar{\rho}_\cdot^{(t,x,h,y)}$ defined by

$$\bar{\rho}_s^{(t,x,h,y)} = x + \frac{s}{t}(2h - x - y), \quad s \in [0,t]. \tag{7.238}$$

Note, that the path $\bar{\rho}_\cdot$ connects x to $2h - y$ during the time interval $[0,t]$. Recall the construction of the functions $\tilde{\phi}_k$ and $\tilde{\phi}_k^\epsilon$ in Theorem 7.3.2.7. The first terms in the representation of $\tilde{\phi}_k$ and $\tilde{\phi}_k^\epsilon$ are integral terms. The integrals range from 0 to $t(h-x)/(2h-x-y)$. The integrands are some particular functions, and integration goes along the path $\rho_\cdot^{(t,x,h,y)}$. Due to what we have just stated, and since the second addends, of which $\tilde{\phi}_k$ or $\tilde{\phi}_k^\epsilon$ consist, neither depend on $\rho_\cdot$ nor on $\bar{\rho}_\cdot$, we are allowed to replace $\rho_\cdot$ in the definition of the functions $\tilde{\phi}_k$ and $\tilde{\phi}_k^\epsilon$ by the path $\bar{\rho}_\cdot$.

But we can also state another interesting fact. Let $k \in \mathbb{N}$. By definition

$$\begin{aligned}
&\phi_k\left(t\frac{(h-y)}{2h-x-y}, h, y\right) \\
&= \int_0^{t\frac{h-y}{2h-x-y}} g_{k-1}\left(t\frac{h-y}{2h-x-y} - s, h + \frac{s}{t\frac{h-y}{2h-x-y}}(y-h), y\right) ds \\
&= \int_0^{t\frac{h-y}{2h-x-y}} g_{k-1}\left(t\frac{h-y}{2h-x-y} - s, h - \frac{s}{t}(2h-x-y), y\right) ds.
\end{aligned}\tag{7.239}$$

If we substitute s by $u - t\frac{h-x}{2h-x-y}$, the integral on the right hand side of the previous equation becomes

$$\begin{aligned}
\phi_k\left(t\frac{(h-y)}{2h-x-y}, h, y\right) &= \int_{t\frac{h-x}{2h-x-y}}^t g_{k-1}\left(t-u, 2h-x-\frac{u}{t}(2h-x-y), y\right) ds \\
&= \int_{t\frac{h-x}{2h-x-y}}^t g_{k-1}\left(t-u, \bar{\bar{\rho}}_u, y\right) du,
\end{aligned}\tag{7.240}$$

where $\bar{\bar{\rho}}_\cdot^{(t,x,h,y)} = \bar{\bar{\rho}}_\cdot$ is the linear path joining $2h-x$ to y during the interval $[0,t]$, defined by

$$\bar{\bar{\rho}}_u = 2h - x - \frac{u}{t}(2h-x-y), \quad u \in [0,t]. \tag{7.241}$$

This means that $\phi_k\left(t\frac{(h-y)}{2h-x-y}, h, y\right)$ is one part of the expression $\phi_k(t, 2h-x, y)$ which is given by

$$\begin{aligned}
\phi_k(t, 2h-x, y) &= \int_0^t g_{k-1}\left(t-u, \bar{\bar{\rho}}_u, y\right) du \\
&= \int_0^{t\frac{h-x}{2h-x-y}} g_{k-1}\left(t-u, \bar{\bar{\rho}}_u, y\right) du + \phi_k\left(t\frac{(h-y)}{2h-x-y}, h, y\right).
\end{aligned}\tag{7.242}$$

By definition, the function $\phi_k(t, 2h-x, y)$ is also a solution to the partial differential equation

$$\frac{\partial}{\partial t}\phi_k\left(t,2h-x,y\right)=\frac{2h-x-y}{t}\frac{\partial}{\partial x}\phi_k\left(t,2h-x,y\right)+g_{n-1}(t,2h-x,y),\tag{7.243}$$

without boundary condition. Consequently, $\phi_k\left(t\frac{(h-y)}{2h-x-y},h,y\right)$ is a part of this solution. For $\phi_k^\epsilon\left(t\frac{(h-y)}{2h-x-y},h,y\right)$ an analogous statement can be made. This property will be useful in Section 7.4.3, where convergence criteria are derived. Our deliberations show that the solutions $\tilde{\phi}_k$ and $\tilde{\phi}_k^\epsilon$ to our above problems with boundary conditions are just linear combinations of parts of solutions of two different problems without boundary conditions, namely those that are constructed from the path $\bar{\rho}.$ defined by (7.238) and from the path $\bar{\bar{\rho}}.$ defined by (7.241), respectively. Finally, note that one obtains $\bar{\bar{\rho}}.$ by reflecting $\bar{\rho}.$ at h. For some more details, also see Appendix 10.3.2.

7.4 Existence of solutions - complete expansions

The aim of this section is to further analyze the functions $\tilde{\phi}_k$ and $\tilde{\phi}_k^\epsilon$ we obtained in the previous section. By reading the Paragraphs 7.3.1 and 7.3.2 thoroughly, one detects that we had to impose existence results and regularity assumptions for the functions $\tilde{\phi}_k$ and $\tilde{\phi}_k^\epsilon$ in order to state our results. No sufficient conditions for existence were given. This shortcoming will be corrected immediately.

In order to avoid misunderstandings let us again define the functions ϕ_k and $\tilde{\phi}_k$ precisely.

Definition 7.4.0.12. Let $T>0$. For $t\in[0,T]$, let $\gamma_\cdot^{(t,x,y)}=\gamma.$ and $\rho_\cdot^{(t,x,h,y)}=\rho.$ denote the paths

$$\gamma_u=x+\frac{u}{t}(y-x),\quad u\in[0,t],\tag{7.244}$$

and

$$\rho_u=\begin{cases}x+\frac{u}{t}(2h-x-y) & ,\quad \text{if }\ 0\le u\le t\frac{h-x}{2h-x-y},\\ y+\frac{t-u}{t}(2h-x-y) & ,\quad \text{if }\ t\frac{h-x}{2h-x-y}<u\le t,\end{cases}\tag{7.245}$$

respectively. For a function $\beta\in\mathcal{C}^\infty(\mathbb{R},\,\mathbb{R})$ we set

$$\phi_1(t,x,y)=-\frac{1}{2}\int_0^t\beta(\gamma_s)ds\quad\text{and}\quad\tilde{\phi}_1(t,x,h,y)=-\frac{1}{2}\int_0^t\beta(\rho_s)ds,\tag{7.246}$$

and then iteratively we define, for $k\ge 2$,

$$\phi_k(t,x,y)=\int_0^t g_{k-1}(t-s,\gamma_s,y)ds\tag{7.247}$$

and

$$\tilde{\phi}_k(t,x,h,y) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{k-1}(t-s,\rho_s,h,y)ds + \phi_k\left(t\frac{h-y}{2h-x-y},h,y\right), \tag{7.248}$$

where

$$g_{k-1}(t,x,y) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{k-1}(t,x,y) - \frac{1}{2}\beta(x)\phi_{k-1}(t,x,y), \tag{7.249}$$

and

$$\tilde{g}_{k-1}(t,x,h,y) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}(t,x,h,y) - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}(t,x,h,y). \tag{7.250}$$

Note again that the second term on the right hand side of (7.248) is the function ϕ_k that stems from the problem without boundary condition evaluated at $(t(h-y)/(2h-x-y),h,y)$.

The first two paragraphs of this section present explicit series expansions with respect to ϵ that solve the differential equations (7.92) and (7.156), respectively, with the correct boundary conditions. The series clearly depend on the function β. Therefore we have to assume that β behaves reasonably. Here, reasonably means that β is such that the respective series converge uniformly. Then, in the last paragraph, we state a result that provides us with a criterion for β to ensure uniform convergence. Overall, this justifies the existence of the results stated in the first two paragraphs. Our statements require tedious calculations. Therefore we moved some of the proofs to Appendix 10.3.

7.4.1 The case without boundary conditions

We begin with an assumption. Below, we will show that this assumption is justified.

Assumption 7.4.1.1. *Let $T>0$ be fixed. Let $\beta \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ and we assume that β has quadratic growth near infinity. Let ϕ_i, $i \in \mathbb{N}$, denote the functions described in Definition 7.4.0.12 above.*
(i) We assume that β is such that the series

$$1 + \sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,y)\in[0,T]\times\bar{S}} \Big|\phi_i(t,x,y)\Big|, \tag{7.251}$$

has a positive radius of convergence with respect to ϵ, for every bounded domain S of $\mathbb{R}^2$. By a domain, we mean an open connected subset of $\mathbb{R}^d$, $d \in \mathbb{N}$. Moreover, we assume that the expression

$$\sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,y)\in[0,T]\times\bar{S}} \Big|\frac{\partial}{\partial t}\phi_i(t,x,y)\Big|, \tag{7.252}$$

and both of the series

$$\sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,y)\in[0,T]\times\bar{S}} \Big|\frac{\partial}{\partial x}\phi_i(t,x,y)\Big| \quad \text{and} \quad \sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,y)\in[0,T]\times\bar{S}} \Big|\frac{\partial^2}{\partial x^2}\phi_i(t,x,y)\Big|, \tag{7.253}$$

have a positive radius of convergence with respect to ϵ for every bounded domain S of $\mathbb{R}^2$. Note that the radius of convergence with respect to the variable ϵ is allowed to depend on T and on the set S.

(ii) Finally, we assume that there is an $\epsilon_0 > 0$ and a positive constant $c > 0$ such that

$$\sum_{i=1}^{\infty} \epsilon^i \sup_{t\in[0,T]} \Big|\frac{\partial}{\partial x}\phi_i(t,x,y)\Big| \leq c \exp\left(c \cdot x^2 \vee y^2\right), \tag{7.254}$$

for all $x, y \in \mathbb{R}$ and for all $0 < \epsilon < \epsilon_0$. The constant c is allowed to depend on T.

Theorem 7.4.1.2. *Let $T > 0$ and let $S \subset \mathbb{R}^2$ be a bounded domain. Let us suppose that the function β satisfies Assumption 7.4.1.1 with a positive radius of convergence ϵ_0. Then the function*

$$v(t,x,y) = 1 + \sum_{i=1}^{\infty} \epsilon^i \phi_i(t,x,y) \tag{7.255}$$

satisfies the differential equation (7.92) *with initial condition* (7.93)*, on $[0,T] \times \bar{S}$ for every $0 < \epsilon < \epsilon_0$. The function*

$$\phi_1^{\epsilon}(t,x,y) = \phi_1(t,x,y) + \sum_{i=1}^{\infty} \epsilon^i \phi_{1+i}(t,x,y) \tag{7.256}$$

satisfies the differential equation (7.99) *with the initial condition* (7.100)*, on $[0,T] \times \bar{S}$ and for every $0 < \epsilon < \epsilon_0$. Moreover, the functions ϕ_n^{ϵ}, $n \in \mathbb{N}$, $n \geq 2$, defined by*

$$\phi_n^{\epsilon}(t,x,y) = \phi_n(t,x,y) + \sum_{i=1}^{\infty} \epsilon^i \phi_{n+i}(t,x,y), \tag{7.257}$$

satisfy the differential equation (7.119) *with initial condition* (7.120)*, on $[0,T] \times \bar{S}$ and for every $0 < \epsilon < \epsilon_0$. Finally,*

$$\lim_{\epsilon\to 0} \phi_n^{\epsilon}(t,x,y) = \phi_n(t,x,y), \quad \forall n \in \mathbb{N}, \tag{7.258}$$

uniformly for (t,x,y) on $[0,T] \times \bar{S}$.

Proof. The proof is straightforward, but it requires some tedious calculations. Hence, it was moved to Appendix 10.3.1. □

Corollary 7.4.1.3. *Let $T > 0$ and $t \in [0,T]$. For $x, y \in \mathbb{R}$, let $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ denote the law of the Brownian bridge, given by* (7.78)*, on $\Omega = \mathcal{C}([0,t],\mathbb{R})$. Suppose that the assumptions* *of Theorem 7.4.1.2 hold. Then, for* $n \in \mathbb{N}$*,*

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\right] = \left\{1 + \sum_{i=1}^{n} \epsilon^i \phi_i(t,x,y) + \epsilon^n \mathcal{R}_\epsilon(t,x,y)\right\}. \tag{7.259}$$

The remainder term satisfies $\lim_{\epsilon\to 0}\mathcal{R}_\epsilon(t,x,y) = 0$*, uniformly for* (t,x,y) *on compact subsets of* $[0,T]\times\mathbb{R}^2$*.*

Proof. First of all, we formally proceed as we did in the proof of Theorem 7.3.1.2. We apply Itô's formula to

$$u \longmapsto \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u, X_u), \quad u \in [0,t]. \tag{7.260}$$

Here, v denotes the function

$$\begin{aligned} v(t,x) \;=\; v(t,x,y) &= 1 + \sum_{i=1}^{\infty} \epsilon^i \phi_i(t,x,y) \\ &= 1 + \sum_{i=1}^{n-1} \epsilon^i \phi_i(t,x,y) + \epsilon^n \phi_n^\epsilon(t,x,y). \end{aligned} \tag{7.261}$$

The result is

$$\begin{aligned} d&\left\{\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u,X_u)\right\} \\ &= -\;\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u,X_u)\frac{\epsilon}{2}\beta(X_u)du \\ &\quad -\;\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{1,0}(t-u,X_u)du \\ &\quad +\;\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u,X_u)dX_u \\ &\quad +\;\frac{1}{2}\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,2}(t-u,X_u)d\langle X\rangle_u. \end{aligned} \tag{7.262}$$

Due to Theorem 7.4.1.2, we know that, for each $y \in \mathbb{R}$, $(t,x) \mapsto v(t,x) = v(t,x,y)$ solves the differential equation (7.119) on compact subsets of $[0,T]\times\mathbb{R}$. If we multiply the previous equation (7.262) with the indicator function $\mathbb{1}_{\{-M<L_t,\,H_t<M\}}$, where $H_t = \sup_{0\le s\le t} X_s$, $L_t = \inf_{0\le s\le t} X_s$ and $M > 0$ denotes a sufficiently large constant, then most of the terms vanish and we obtain

$$\begin{aligned} d&\left\{\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v(t-u,X_u)\right\}\mathbb{1}_{\{-M<L_t,\,H_t<M\}} \\ &= \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u,X_u)\sqrt{\epsilon}dB_u\,\mathbb{1}_{\{-M<L_t,\,H_t<M\}}. \end{aligned} \tag{7.263}$$

Assumption 7.4.1.1 *(ii)*, in combination with the fact that β was assumed to satisfy a quadratic growth condition, implies that the expression

$$\exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u) \tag{7.264}$$

is square-integrable if ϵ is sufficiently small. Hence, we are allowed to perform the following calculation:

$$\begin{aligned}
0 &= \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u)\sqrt{\epsilon}dB_u\right] \\
&= \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u)\sqrt{\epsilon}dB_u \mathbb{1}_{\{-M<L_t,\, H_t<M\}}\right] \\
&+ \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u)\sqrt{\epsilon}dB_u \mathbb{1}_{\{-M<L_t,\, H_t<M\}^c}\right].
\end{aligned} \tag{7.265}$$

This means that, in order to bound the right hand side of (7.263), we have to estimate the term

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u)\sqrt{\epsilon}dB_u \mathbb{1}_{\{-M<L_t, H_t<M\}^c}\right]. \tag{7.266}$$

Without loss of generality we assume that $-\beta$ is positive, and, as a first step, we estimate the expression

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u)\sqrt{\epsilon}dB_u \mathbb{1}_{\{H_t\geq M\}}\right]. \tag{7.267}$$

Due to Assumption 7.4.1.1 *(ii)* and since β has quadratic growth, we obtain

$$\begin{aligned}
&\left|\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\int_0^t \exp\left(-\frac{\epsilon}{2}\int_0^u \beta(X_s)ds\right) v_{0,1}(t-u, X_u)\sqrt{\epsilon}dB_u \mathbb{1}_{\{H_t\geq M\}}\right]\right| \\
&\leq \sqrt{tc\epsilon\, \hat{\mathbb{E}}_x^{y,\epsilon,t}\Big[\exp\left(tc\epsilon H_t^2\right)\exp\left(2c\cdot H_t^2 \vee y^2\right)\Big]}\sqrt{\hat{\mathbb{P}}_x^{y,\epsilon,t}[H_t\geq M]},
\end{aligned} \tag{7.268}$$

with a sufficiently large, positive constant c. Consequently, if ϵ is sufficiently small, the expectation on the right hand side of (7.268) exists. This follows directly from the fact that, for $a \geq x \vee y$,

$$\hat{\mathbb{P}}_x^{y,\epsilon,t}\left[H_t \geq a\right] = \exp\left(-2\frac{(a-x)(a-y)}{\epsilon t}\right). \tag{7.269}$$

This expression also permits to estimate the second term on the right hand side of (7.268). The expression (7.266) can now easily be bounded in an analogous way by a distinction of cases. By taking expectations on both sides of (7.263), we obtain the intermediate result

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\mathbb{1}_{\{-M<L_t,\, H_t<M\}}\right]$$

$$= \; 1 + \sum_{i=1}^{n-1} \epsilon^i \phi_i(t,x,y) + \epsilon^n \phi_n^\epsilon(t,x,y) + r_{exp}(\epsilon,t,x,y), \tag{7.270}$$

where $r_{exp}(\epsilon,t,x,y)$ is a term that converges exponentially fast to 0 as $\epsilon \to 0$. By the above estimates, convergence is uniform for (t,x,y) on compact subsets of $[0,T] \times \mathbb{R}^2$. An estimate similar to (7.268) shows that the left hand side of (7.270) deviates from

$$\hat{\mathbb{E}}_x^{y,\epsilon,t} \left[\exp\left(-\frac{\epsilon}{2} \int_0^t \beta(X_u)du \right) \right] \tag{7.271}$$

only by quantity that is also exponentially negligible when $\epsilon \to 0$. The proof now follows immediately, since $\lim_{\epsilon\to 0} \phi_n^\epsilon = \phi_n$, uniformly on compact subsets of $[0,T] \times \mathbb{R}^2$. $\square$

7.4.2 The case with boundary conditions

We turn our attention to the case with boundary condition. Again we make an assumption. In Paragraph 7.4.3 we will then show that this assumption is justified.

Assumption 7.4.2.1. *Let $T > 0$ be fixed. Let $\beta \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ and we assume that β has quadratic growth near infinity. Let us consider the functions ϕ_i and $\tilde{\phi}_i$ of Definition 7.4.0.12.*

(i) We assume that β is such that the series

$$1 + \sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,h,y)\in[0,T]\times\bar{S}} \left| \tilde{\phi}_i(t,x,h,y) \right| \tag{7.272}$$

has a positive radius of convergence with respect to ϵ for every bounded domain $S \subset \{(x,h,y) \in \mathbb{R}^3 \,|\, x,y \le h\}$. Moreover, we assume that the expression

$$\sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,h,y)\in[0,T]\times\bar{S}} \left| \frac{\partial}{\partial t} \tilde{\phi}_i(t,x,h,y) \right| \tag{7.273}$$

and both of the series

$$\sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,h,y)\in[0,T]\times\bar{S}} \left| \frac{\partial}{\partial x} \tilde{\phi}_i(t,x,h,y) \right| \quad \text{and} \quad \sum_{i=1}^{\infty} \epsilon^i \sup_{(t,x,h,y)\in[0,T]\times\bar{S}} \left| \frac{\partial^2}{\partial x^2} \tilde{\phi}_i(t,x,h,y) \right| \tag{7.274}$$

have a positive radius of convergence radius with respect to ϵ for every bounded domain $S \subset \{(x,h,y) \in \mathbb{R}^3 \,|\, x,y \le h\}$. Note that the radius of convergence with respect to the variable ϵ is allowed to depend on T and on the set S.

(ii) Finally, we assume that there is an $\epsilon_0 > 0$ and a positive constant $c > 0$ such that

$$\sum_{i=1}^{\infty} \epsilon^i \sup_{t\in[0,T]} \left| \frac{\partial}{\partial x} \tilde{\phi}_i(t,x,h,y) \right| \le c \exp\left(c \cdot |2h-x|^2 \vee |y|^2 \vee |2h-y|^2 \vee |x|^2 \right), \tag{7.275}$$

for all $x, h, y \in \mathbb{R}$, with $x, y \le h$, and for all $0 < \epsilon < \epsilon_0$. The constant c is allowed to depend on T.

Theorem 7.4.2.2. *Assume that for $T > 0$ and for a bounded domain $S \subset \{(x,h,y) \in \mathbb{R}^3 \,|\, x, y \le h\}$ the function $\beta \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ satisfies both Assumption 7.4.1.1 and Assumption 7.4.2.1 with a positive radius of convergence $\epsilon_0 > 0$.*

(i) The function $\tilde{\phi}_1^\epsilon$, given by

$$\tilde{\phi}_1^\epsilon(t,x,h,y) = \tilde{\phi}_1(t,x,h,y) + \sum_{i=1}^{\infty} \epsilon^i \tilde{\phi}_{1+i}(t,x,h,y) = \sum_{i=1}^{\infty} \epsilon^{i-1} \tilde{\phi}_i(t,x,h,y), \quad (7.276)$$

satisfies, for every $0 < \epsilon < \epsilon_0$ and on the set $[0,T] \times \bar{S}$, the differential equation (7.166) with initial condition (7.167) and with boundary condition (7.168). Moreover, for every $0 < \epsilon < \epsilon_0$ and on $[0,T] \times \bar{S}$, the functions $\tilde{\phi}_k^\epsilon$, $k \in \mathbb{N}$, defined by

$$\tilde{\phi}_k^\epsilon(t,x,h,y) = \tilde{\phi}_k(t,x,h,y) + \sum_{i=1}^{\infty} \epsilon^i \tilde{\phi}_{k+i}(t,x,h,y) = \sum_{i=k}^{\infty} \epsilon^{i-k} \tilde{\phi}_i(t,x,h,y), \quad (7.277)$$

satisfy the differential equation (7.199) with initial condition (7.200) and boundary condition (7.201). Finally,

$$\lim_{\epsilon \to 0} \tilde{\phi}_k^\epsilon(t,x,h,y) = \tilde{\phi}_k(t,x,h,y), \quad (7.278)$$

uniformly for (t,x,h,y) on $[0,T] \times \bar{S}$.

(ii) Furthermore, on the set $[0,T] \times \bar{S}$, the function v_h, which is defined by

$$v_h(t,x,h,y) = v(t,x,y) - \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left\{1 + \sum_{i=1}^{\infty} \epsilon^i \tilde{\phi}_i(t,x,h,y)\right\}, \quad (7.279)$$

with v given by (7.255), satisfies, for every $0 < \epsilon < \epsilon_0$, the differential equation (7.156) with initial condition (7.157) and boundary condition (7.158).

Proof. The proof is straightforward. However, it requires some tedious calculations. It can be found in Appendix 10.3.1. □

Corollary 7.4.2.3. *Let $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ denote the law of (7.78) on $\Omega = \mathcal{C}([0,t],\mathbb{R})$. Assume that the assumptions of Theorem 7.4.2.2 hold, then, for $n \in \mathbb{N}$,*

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\mathbb{1}_{\{\sup_{0\le u\le t} X_u \ge h\}}\right]$$
$$= \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left\{1 + \sum_{i=1}^{n} \epsilon^i \tilde{\phi}_i(t,x,h,y) + \epsilon^n \mathcal{R}_\epsilon(t,x,h,y)\right\}. \quad (7.280)$$

The remainder term $\mathcal{R}_\epsilon(t,x,h,y)$ *converges to* 0 *as* $\epsilon \to 0$. *Convergence is uniform for* (t,x,h,y) *on compact subsets of* $[0,T] \times \{(x,h,y) \in \mathbb{R}^3 \,|\, x, y \le h\}$.

Proof. Let $n \in \mathbb{N}$. From the proof of Corollary 7.4.1.3 we know that

$$\begin{aligned}\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\right] &= 1 + \sum_{i=1}^{n-1}\epsilon^i \phi_i(t,x,y) + \epsilon^n \phi_n^\epsilon(t,x,y) + r_{exp}(\epsilon,t,x,y)\\ &= v(t,x,y) + r_{exp}(\epsilon,t,x,y), \end{aligned} \tag{7.281}$$

where $r_{exp}(\epsilon,t,x,y)$ is exponentially negligible as $\epsilon \to 0$. By mimicking the procedure in the proof of Corollary 7.4.1.3 and by combining it with the results of Theorem 7.4.2.2 *(ii)*, it is straightforward to show that there is another function $\widetilde{r}_{exp}(\epsilon,t,x,h,y)$, such that

$$\begin{aligned}&\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\mathbb{1}_{\left\{\sup_{0\le u\le t} X_u < h\right\}}\right]\\ &= \hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\right]\\ &\quad - \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\Big\{1 + \sum_{i=1}^{n-1}\epsilon^i\tilde{\phi}_i(t,x,h,y) + \epsilon^n\tilde{\phi}_n^\epsilon(t,x,h,y)\\ &\qquad + \widetilde{r}_{exp}(\epsilon,t,x,h,y)\Big\}. \end{aligned} \tag{7.282}$$

The function $\widetilde{r}_{exp}(\epsilon,t,x,h,y)$ converges to 0 exponentially fast as $\epsilon \to 0$, and convergence is uniform on compact subsets of $[0,T]\times\{(x,h,y) \in \mathbb{R}^3 \,|\, x,y \le h\}$. The proof now follows immediately, since $\lim_{\epsilon\to 0}\tilde{\phi}_n^\epsilon = \tilde{\phi}_n$, uniformly on compact subsets of $[0,T] \times \{(x,h,y) \in \mathbb{R}^3 \,|\, x,y \le h\}$. □

7.4.3 Conditions for uniform convergence

In the previous paragraph, we found series expansions of

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\right] \tag{7.283}$$

and

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\mathbb{1}_{\{\sup_{0\le u\le t} X_u \ge h\}}\right], \tag{7.284}$$

with respect to ϵ. But we had to postulate the convergence of the series. In the present section, we present a result which states uniform convergence in an important case, namely, if β is a quadratic polynomial. For convenience we moved some of our calculations to Appendix 10.3. We first show that Assumption 7.4.1.1 is satisfied for quadratic polynomials.

The case without boundary condition

Proposition 7.4.3.1. *Assume that $\beta : \mathbb{R} \to \mathbb{R}$ is a quadratic polynomial. Let the functions ϕ_k be as described in Definition 7.4.0.12. Then the series*

$$1 + \sum_{i=1}^{\infty} \epsilon^i \phi_i(t,x,y) \tag{7.285}$$

satisfies Assumption 7.4.1.1.

Proof. In Proposition 10.3.2.1 in Appendix 10.3.2 we derive upper bounds for the functions ϕ_k if $\beta(x) = x^2$ or $-x^2$. They are given by

$$\Big|\phi_k(t,x,y)\Big| = 2^k t^k \sum_{j=1}^{2^k} \frac{(|x|^2 \vee |y|^2)^j}{j!!}. \tag{7.286}$$

This particularly implies the estimate

$$\Big|\phi_k(t,x,y)\Big| \le 2^k t^k \exp\Big(|x|^2 \vee |y|^2\Big), \tag{7.287}$$

which in turn implies that

$$\sum_{k=0}^{\infty} \epsilon^k \Big|\phi_k(t,x,y)\Big| \le \exp\Big(|x|^2 \vee |y|^2\Big) \cdot \sum_{k=0}^{\infty} \epsilon^k t^k 2^k. \tag{7.288}$$

For any $x, y \in \mathbb{R}$, the series on the right hand side converges if $\epsilon < \frac{1}{2t}$. And, due to the estimate (7.288), convergence is uniform for (x,y) on compact subsets of $\mathbb{R}^2$. By going carefully through the proofs of Lemma 10.3.2.2 and Lemma 10.3.2.3 one can show that the following bounds hold:

$$\sum_{k=0}^{\infty} \epsilon^k \Big|\frac{\partial}{\partial t}\phi_k(t,x,y)\Big| \le \exp\Big(|x|^2 \vee |y|^2\Big) \cdot \sum_{k=1}^{\infty} \epsilon^k k t^{k-1} 2^k, \tag{7.289}$$

$$\sum_{k=0}^{\infty} \epsilon^k \Big|\frac{\partial}{\partial x}\phi_k(t,x,y)\Big| \le (|x| \vee |y|) \exp\Big(|x|^2 \vee |y|^2\Big) \cdot \sum_{k=1}^{\infty} \epsilon^k t^k 2^k, \tag{7.290}$$

$$\sum_{k=0}^{\infty} \epsilon^k \Big|\frac{\partial^2}{\partial x^2}\phi_k(t,x,y)\Big| \le (|x|^2 \vee |y|^2) \exp\Big(|x|^2 \vee |y|^2\Big) \cdot \sum_{k=2}^{\infty} \epsilon^k (k-1) t^k 2^k. \tag{7.291}$$

The same reasoning as above shows that for each series there is a positive radius of convergence, and thereby Assumption 7.4.1.1 holds for $\beta(x) = x^2$ and $\beta(x) = -x^2$. Assumption 7.4.1.1 *(ii)* is trivial to check. For the linear functions $\beta(x) = \pm x$, the estimates of ϕ_k basically remain the same. This is even easier to prove than for the case $\beta(x) = x^2$. Thus, for a more general quadratic polynomial $\beta(x) = c_2 x^2 + c_1 x + c_0$ with $c_0, c_1, c_2 \in \mathbb{R}$, we are able to find estimates similar to (7.288) and (7.289)-(7.291). The only difference is that in this general case the estimates of ϕ_k and its derivatives will depend on $|c_2|\,|x|^2 + |c_1|\,|x| + |c_0|$ and $|c_2|\,|y|^2 + |c_1|\,|y| + |c_0|$. We omit further details here, but Assumption 7.4.1.1 is clearly satisfied for any quadratic polynomial with real valued coefficients. $\square$

Let us now move on to the study of Assumption 7.4.2.1.

The case with boundary condition

Proposition 7.4.3.2. *Assume that $\beta : \mathbb{R} \to \mathbb{R}$ is a quadratic polynomial. Let the functions ϕ_k and $\tilde{\phi}_k$ be as described in Definition 7.4.0.12. Then the series*

$$1 + \sum_{i=1}^{\infty} \epsilon^i \phi_i(t,x,y) \quad \text{and} \quad 1 + \sum_{i=1}^{\infty} \epsilon^i \tilde{\phi}_i(t,x,h,y) \tag{7.292}$$

satisfy Assumption 7.4.1.1 and Assumption 7.4.2.1.

Proof. In Proposition 10.3.2.4 in Appendix 10.3.2, upper bounds are derived for the functions $\tilde{\phi}_k$ for the case where $\beta(x) = \pm\, x^2$. These bounds are given by

$$\left|\tilde{\phi}_k(t,x,h,y)\right| = 10^k t^k \exp\left(\left\{|2h-x|^2 \vee |y|^2\right\} + \left\{|2h-y|^2 \vee |x|^2\right\}\right), \tag{7.293}$$

which implies that

$$\sum_{k=0}^{\infty} \epsilon^k \left|\tilde{\phi}_k(t,x,y)\right| \le \exp\left(\left\{|2h-x|^2 \vee |y|^2\right\} + \left\{|2h-y|^2 \vee |x|^2\right\}\right) \cdot \sum_{k=0}^{\infty} \epsilon^k 10^k t^k. \tag{7.294}$$

For any $x, y \in \mathbb{R}$, the series on the right hand side converges if $\epsilon < \frac{1}{10t}$. And, due to the estimate (7.294), convergence is uniform for (x,h,y) on compact subsets of $\{(x,h,y) \in \mathbb{R}^3 \,|\, x, y \le h\}$.

By going carefully through the proofs of Lemma 10.3.2.5 and Lemma 10.3.2.8 one can show that the following estimate holds:

$$\begin{aligned} &\sum_{k=0}^{\infty} \epsilon^k \Big|\frac{\partial}{\partial t} \tilde{\phi}_k(t,x,h,y)\Big| \\ &\le \exp\left(\left\{|2h-x|^2 \vee |y|^2\right\} + \left\{|2h-y|^2 \vee |x|^2\right\}\right) \cdot \sum_{k=1}^{\infty} \epsilon^k k t^{k-1} 10^k. \end{aligned} \tag{7.295}$$

We are equally able to find bounds for the first derivatives of $\tilde{\phi}_k(t,x,h,y)$ with respect to x. The resulting upper bound for the series is given by

$$\begin{aligned} &\sum_{k=0}^{\infty} \epsilon^k \Big|\frac{\partial}{\partial x} \tilde{\phi}_k(t,x,h,y)\Big| \\ &\le \left\{|2h-x| \vee |y|\right\} \exp\left(\left\{|2h-x|^2 \vee |y|^2\right\} + \left\{|2h-y|^2 \vee |x|^2\right\}\right) \cdot \sum_{k=1}^{\infty} \epsilon^k t^k 10^k \end{aligned}$$

$$+ \ \Big\{|2h-y| \vee |x|\Big\} \exp\Big(\Big\{|2h-x|^2 \vee |y|^2\Big\} + \Big\{|2h-y|^2 \vee |x|^2\Big\}\Big) \cdot \sum_{k=1}^{\infty} \epsilon^k t^k 10^k. \tag{7.296}$$

For the series consisting of second derivatives of $\tilde{\phi}_k(t,x,h,y)$ with respect to x, we obtain the upper bound

$$\begin{aligned}
&\sum_{k=0}^{\infty} \epsilon^k \Big| \frac{\partial^2}{\partial x^2} \tilde{\phi}_k(t,x,h,y) \Big| \\
&\le \Big\{|2h-x|^2 \vee |y|^2\Big\} \exp\Big(\Big\{|2h-x|^2 \vee |y|^2\Big\} + \Big\{|2h-y|^2 \vee |x|^2\Big\}\Big) \cdot \sum_{k=1}^{\infty} \epsilon^k (k-1) t^k 10^k \\
&+ \ 2\Big\{|2h-x| \vee |y|\Big\} \cdot \Big\{|2h-y|^2 \vee |x|^2\Big\} \exp\Big(\Big\{|2h-x|^2 \vee |y|^2\Big\} + \Big\{|2h-y|^2 \vee |x|^2\Big\}\Big) \cdot \sum_{k=1}^{\infty} \epsilon^k t^k 10^k 2^k \\
&+ \ \Big\{|2h-y|^2 \vee |x|^2\Big\} \exp\Big(\Big\{|2h-x|^2 \vee |y|^2\Big\} + \Big\{|2h-y|^2 \vee |x|^2\Big\}\Big) \cdot \sum_{k=1}^{\infty} \epsilon^k (k-1) t^k 10^k.
\end{aligned} \tag{7.297}$$

The same reasoning as above shows that for each of the series (7.295)-(7.297) there is a positive radius of convergence and hence, Assumption 7.4.2.1 holds for $\beta(x) = x^2$ and $\beta(x) = -x^2$. It is trivial to see that Assumption 7.4.2.1 *(ii)* is satisfied. Again there is no limitation in the sense that we are also able to find similar bounds for $\tilde{\phi}_k$ and its derivatives if β is a general quadratic polynomial. Thus, we have shown that Assumption 7.4.2.1 holds for quadratic polynomials with real valued coefficients. □

Assessment of the convergence criteria

The validity of Theorem 7.4.1.2 and Theorem 7.4.2.2 is justified by the Propositions 7.4.3.1 and 7.4.3.2, in hindsight. Let us note that other criteria are imaginable to imply uniform convergence. For every candidate $\beta : \mathbb{R} \to \mathbb{R}$ one has to verify if the series (7.255) and (7.279) are uniformly convergent – either on the whole space $\mathbb{R}_+ \times \mathbb{R}^3$ or at least on compact subsets. Then the results of the present Paragraph 7.4 yield approximations to the quantities (7.283) and (7.284). We have reason to believe that there is a class of functions $\beta \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ with quadratic growth near infinity that satisfy the necessary convergence conditions. Furthermore, a function $\beta \in \mathcal{C}_b^\infty(\mathbb{R},\mathbb{R})$ with uniformly bounded derivatives of all orders most likely also satisfies the convergence criteria. However, the proof in these cases is more difficult. A concrete classification of the functions $\beta : \mathbb{R} \to \mathbb{R}$ for which Assumption 7.4.1.1 and Assumption 7.4.2.1 are satisfied remains to be determined. This gives rise to further research.

Finally, we stress that general polynomials of degree > 2 presumably do not satisfy Assumption 7.4.1.1 or Assumption 7.4.1.1. Let us explain, why we have this opinion. If the series

$$1 + \sum_{i=1}^{n} \epsilon^i \phi_i(t,x,y) \tag{7.298}$$

converges, it coincides with the expression

$$\hat{\mathbb{E}}_x^{y,\epsilon,t}\left[\exp\left(-\frac{\epsilon}{2}\int_0^t \beta(X_u)du\right)\right]. \tag{7.299}$$

Under the measure $\hat{\mathbb{P}}_x^{y,\epsilon,t}$, the coordinate process X evolves as a Brownian bridge. But quadratic forms of Gaussian random variables have at most exponential moment. Therefore, we cannot expect much for the case $\beta(x) = x^k$ with $|k| > 2$. However, the expectation (7.299) always exists if β is strictly positive. For example, this is the case if $\beta(x) = x^{2k}$ with $k \geq 1$. In this case it might be possible to find estimates for ϕ_k and $\tilde{\phi}_k$ and convergence results for the respective series that guarantee the compliance with Assumption 7.4.1.1 and Assumption 7.4.1.1. In order to show such a result, it will be necessary to take the signs of the different terms of which the coefficients ϕ_k and $\tilde{\phi}_k$ consist into account. This certainly will cause a lot of technical difficulties.

7.5 Extension of the results for pinned diffusion

We are now going to extend the results of Proposition 7.2.1.2 for the diffusion model (7.3), where we still assume that the diffusion coefficient satisfies $\sigma \equiv 1$. We use the same notations as in Paragraph 7.2.1. Particularly, recall formulae (7.21), (7.22) and (7.23). Moreover, let the functions ϕ_k and $\tilde{\phi}_k$ be recursively defined according to Definition 7.4.0.12. For $\epsilon > 0$ and $n \in \mathbb{N}$, let $\Phi_\epsilon^{(n)}$ be defined by

$$\frac{p_\epsilon(1,x,y)e^{G(y)-G(x)}}{q_\epsilon(1,x,y)} = 1 + \sum_{k=1}^{n-1} \epsilon^k \Phi^{(k)}(x,y) + \epsilon^n \Phi_\epsilon^{(n)}(x,y), \tag{7.300}$$

where the expressions $\Phi^{(k)}(x,y),\ k \in \mathbb{N},\ k < n$, denote the functions

$$\Phi^{(k)}(x,y) = \sum_{j=0}^{k}(-1)^j \sum_{(z_1,\ldots,z_k)\in M_{j,k}} \binom{j}{z_1,\ldots,z_k} \phi_1(1,x,y)^{z_1}\cdot\ldots\cdot\phi_k(1,x,y)^{z_k}. \tag{7.301}$$

The set $M_{j,k}$, in the previous formula, is defined by

$$M_{j,k} = \left\{(z_1,\ldots,z_k) \in \mathbb{N}_0^k \middle| \sum_{\nu=1}^{k} z_\nu = j,\ \sum_{\nu=1}^{k} \nu\cdot z_\nu = k\right\}. \tag{7.302}$$

Moreover, let $A_h = A_h(x,y) \in \mathcal{F}_1$ be the set

$$A_h(x,y) = \left\{ g \in \mathcal{C}^1([0,1],\mathbb{R}) \,\middle|\, g(0) = x,\ g(1) = y,\ \sup_{0\leq u\leq 1} g(u) \geq h\right\}, \tag{7.303}$$

and let $\widetilde{\Phi}_\epsilon^{(n)}$ be defined by the equation

$$\begin{aligned}\hat{\mathbb{E}}_x^{y,\epsilon}&\left[\exp\left(-\frac{\epsilon}{2}\int_0^1[\mu'(X_u)+\mu(X_u)^2]du\right)\mathbb{1}_{\{X\in A_h\}}\right]\\ &=\hat{\mathbb{P}}_x^{y,\epsilon}[X\in A_h]\left\{1+\sum_{k=1}^{n-1}\epsilon^k\widetilde{\Phi}^{(k)}(x,h,y)+\epsilon^n\widetilde{\Phi}_\epsilon^{(n)}(x,h,y)\right\},\end{aligned}\tag{7.304}$$

where

$$\widetilde{\Phi}^{(k)}(x,h,y)=\tilde{\phi}_k(1,x,h,y),\quad k=1,\dots,n-1.\tag{7.305}$$

The following proposition holds.

Proposition 7.5.0.3. *Let the coefficient μ be infinitely many times differentiable and, for $\beta=\mu'+\mu^2$, consider the functions ϕ_k and $\tilde{\phi}_k$ of Definition 7.4.0.12. Moreover, we assume that β satisfies the Assumptions 7.4.1.1 and 7.4.2.1. For $n\in\mathbb{N}$, the functions $\Phi_\epsilon^{(n)}(x,y)$ and $\widetilde{\Phi}_\epsilon^{(n)}(x,h,y)$, defined by* (7.300) *and* (7.304)*, respectively, satisfy the following two properties.*

(i) If $M_{j,n}$ denotes the set (7.302)*, then*

$$\begin{aligned}&\lim_{\epsilon\to0}\Phi_\epsilon^{(n)}(x,y)\\ &=\sum_{j=0}^n(-1)^j\sum_{(z_1,\dots,z_n)\in M_{j,n}}\binom{j}{z_1,\dots,z_n}\phi_1(1,x,y)^{z_1}\cdot\ldots\cdot\phi_n(1,x,y)^{z_n},\end{aligned}\tag{7.306}$$

uniformly for (x,y) on compact subsets of $\mathbb{R}^2$.

(ii) Let $A_h=A_h(x,y)\in\mathcal{F}_1$ denote the set of paths (7.303)*. Then the function $\widetilde{\Phi}_\epsilon^{(n)}$, defined by the equation* (7.304)*, satisfies*

$$\lim_{\epsilon\to0}\widetilde{\Phi}_\epsilon^{(n)}(x,h,y)=\tilde{\phi}_n(1,x,h,y),\tag{7.307}$$

uniformly for (x,h,y) on compact subsets of $\{(x,h,y)\in\mathbb{R}^3\,|\,x,y\le h\}$.

Proof. (i) If $n=1$ the above assertion is exactly the same as the one stated in Proposition 7.2.1.2. In the general case, if $\ell(A)$ denotes the Lebesgue measure of the Borel set $A\in\mathcal{B}(\mathbb{R})$, it is obvious that

$$p_\epsilon(1,x,y)=\lim_{\delta\to0}\frac{\mathbb{P}_x^\epsilon[X_1\in U_\delta(y)]}{\ell(U_\delta(y))}\tag{7.308}$$

and

$$q_\epsilon(1,x,y)=\lim_{\delta\to0}\frac{\mathbb{Q}_x^\epsilon[X_1\in U_\delta(y)]}{\ell(U_\delta(y))}$$

$$= \lim_{\delta \to 0} \frac{\mathbb{E}_x^\epsilon \left[e^{G(X_1) - G(x) - (\epsilon/2) \int_0^1 \beta(X_u) du} \mathbb{1}_{\{X_1 \in U_\delta(y)\}} \right]}{\ell(U_\delta(y))}. \tag{7.309}$$

Here $U_\delta(y)$ denotes the ball having radius δ and being centered around y. By definition, we have

$$1 + \sum_{k=1}^{n-1} \epsilon^k \Phi^{(k)}(x, y) + \epsilon^n \Phi_\epsilon^{(n)}(x, y)$$

$$= \lim_{\delta \to 0} \frac{1}{\mathbb{E}_x^\epsilon \left[e^{G(X_1) - G(y) - (\epsilon/2) \int_0^1 \beta(X_u) du} \Big| X_1 \in U_\delta(y) \right]}. \tag{7.310}$$

We must show that

$$\lim_{\epsilon \to 0} \Phi_\epsilon^{(n)}(x, y) = \sum_{j=1}^{n} (-1)^j \sum_{(z_1, \dots, z_n) \in M_{j,n}} \binom{j}{z_1, \dots, z_n} \phi_1(1, x, y)^{z_1} \cdot \ldots \cdot \phi_n(1, x, y)^{z_n}, \tag{7.311}$$

where $M_{j,n}$ is the set defined by (7.302). The proof will be conducted in two steps.

First step. Recall, that the measure $\mathbb{P}_x^\epsilon$ denotes the law of the Brownian motion $(x + \sqrt{\epsilon} B_s,\ s \geq 0)$, see also (7.11), and that $\hat{\mathbb{P}}_x^{y,\epsilon}$ denotes the law of the Brownian bridge defined by (7.78). Initially, we concentrate on the analysis of the two terms

$$\mathbb{E}_x^\epsilon \left[e^{G(X_1) - G(y) - (\epsilon/2) \int_0^1 \beta(X_u) du} \Big| X_1 \in U_\delta(y) \right] \tag{7.312}$$

and

$$\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2) \int_0^1 \beta(X_u) du} \right]. \tag{7.313}$$

Assumption 7.4.1.1 implies that the series defining (7.313) converges uniformly on compact subsets of $\mathbb{R}^2$. Compare the statement of Corollary 7.4.1.3. Thus, there exists an $\epsilon_0 > 0$, such that (7.313) is uniformly bounded for (x, y) on compact subsets of $\mathbb{R}^2$ and for all $0 < \epsilon < \epsilon_0$. The function G is continuous by definition. Consequently, the expression (7.312) is also uniformly bounded for (x, y) on compact subsets of $\mathbb{R}^2$ as $\delta \to 0$, and for all $0 < \epsilon < \epsilon_0$.

We intend to make use of the following property:

$$\mathbb{P}_x^\epsilon \left[\cdot \,\big|\, X_1 \in U_\delta(y) \right] \longrightarrow \hat{\mathbb{P}}_x^{y,\epsilon}, \tag{7.314}$$

weakly as $\delta \to 0$. A proof for this fact is given in Billingsley [15], page 101 ff.

One encounters difficulties if β is unbounded. But the asymptotics for $\epsilon \to 0$ of the two expressions (7.312) and (7.313) only change by a quantity that is exponentially negligible if β is changed outside a compact interval that contains x and y. To see this, let $c > 0$ be a sufficiently large constant. Without loss of generality, we can assume that $-\beta$ is a positive function. The general case follows directly by a distinction of cases. We obtain the estimate

$$\begin{aligned}
&\left|\mathbb{E}_x^{\epsilon}\left[e^{G(X_1)-G(y)-(\epsilon/2)\int_0^1 \beta(X_u)du}\Big|X_1 \in U_\delta(y)\right] - \hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du}\right]\right| \\
&\le \left|\mathbb{E}_x^{\epsilon}\left[e^{G(X_1)-G(y)-(\epsilon/2)\int_0^1 \beta(X_u)du}\mathbb{1}_{\{H_1<c\}}\Big|X_1 \in U_\delta(y)\right] - \hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du}\mathbb{1}_{\{H_1<c\}}\right]\right| \\
&\quad + \mathbb{E}_x^{\epsilon}\left[e^{G(X_1)-G(y)-(\epsilon/2)\int_0^1 \beta(X_u)du}\mathbb{1}_{\{H_1\ge c\}}\Big|X_1 \in U_\delta(y)\right] \\
&\quad + \hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du}\mathbb{1}_{\{H_1\ge c\}}\right],
\end{aligned} \tag{7.315}$$

where $H_1 = \sup_{0\le s\le 1} X_s$. By the continuity of G, it is straightforward to show that the first term on the right hand side of inequality (7.315) converges to 0 as $\delta \to 0$. Moreover, for every $\delta_0 > 0$, there is a constant K (that is allowed to depend on y), such that, for all $0 < \delta < \delta_0$,

$$\begin{aligned}
&\mathbb{E}_x^{\epsilon}\left[e^{G(X_1)-G(y)-(\epsilon/2)\int_0^1 \beta(X_u)du}\mathbb{1}_{\{H_1\ge c\}}\Big|X_1 \in U_\delta(y)\right] \\
&\qquad \le K\sqrt{\mathbb{E}_x^{\epsilon}\left[e^{-\epsilon\int_0^1 \beta(X_u)du}\Big|X_1 \in U_\delta(y)\right]}\sqrt{\mathbb{P}_x^{\epsilon}\left[H_1 \ge c\Big|X_1 \in U_\delta(y)\right]}.
\end{aligned} \tag{7.316}$$

For the last term in (7.315), we have the following estimate

$$\hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du}\mathbb{1}_{\{H_1\ge c\}}\right] \le \sqrt{\hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-\epsilon\int_0^1 \beta(X_u)du}\right]}\sqrt{\hat{\mathbb{P}}_x^{y,\epsilon}\left[H_1 \ge c\right]}. \tag{7.317}$$

Let us note that

$$\lim_{\delta\to 0}\mathbb{P}_x^{\epsilon}\left[H_1 \ge c\Big|X_1 \in U_\delta(y)\right] = \hat{\mathbb{P}}_x^{y,\epsilon}\left[H_1 \ge c\right] = \exp\left(-2\frac{(c-x)(c-y)}{\epsilon}\right). \tag{7.318}$$

The previous deliberations, in combination with the boundedness of the expressions (7.312) and (7.313), show that

$$\begin{aligned}
&\lim_{\delta\to 0}\mathbb{E}_x^{\epsilon}\left[e^{G(X_1)-G(y)-(\epsilon/2)\int_0^1 \beta(X_u)du}\Big|X_1 \in U_\delta(y)\right] \\
&= \hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du}\right] + \widehat{r}_{exp}(\epsilon, x, y),
\end{aligned} \tag{7.319}$$

where the term $\widehat{r}_{exp}(\epsilon, x, y)$ converges exponentially fast to 0 as $\epsilon \to 0$. By our estimates, convergence is uniform for (x, y) on compact subsets of $\mathbb{R}^2$. Since the right hand side of

(7.319) is strictly positive for sufficiently small ϵ, we obtain the intermediate result

$$\lim_{\delta\to 0}\frac{1}{\mathbb{E}_x^{\epsilon}\left[e^{G(X_1)-G(y)-(\epsilon/2)\int_0^1\beta(X_u)du}\big|X_1\in U_\delta(y)\right]}=\frac{1}{\hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1\beta(X_u)du}\right]+\widehat{r}_{exp}(\epsilon,x,y)}. \tag{7.320}$$

Due to our assumptions, the expectation $\hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1\beta(X_u)du}\right]$ in the denominator of (7.320) converges to 1, uniformly for compact subsets of $\mathbb{R}^2$ as ϵ tends to 0. Let us stress again that the term $\widehat{r}_{exp}$ is exponentially negligible for (x,y) on compact subsets of $\mathbb{R}^2$ as $\epsilon\to 0$. Consequently, this extra term does not affect our analysis of the asymptotic behavior. It can simply be omitted in the sequel.

Second step. We are now going to derive an expansion of the term

$$\frac{1}{\hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1\beta(X_u)du}\right]}. \tag{7.321}$$

The idea is to make use of the following series expansion: for $z\in(0,\infty)$ and $n\in\mathbb{N}$,

$$\frac{1}{z}=\sum_{k=0}^{n}(-1)^k(z-1)^k+(-1)^{n+1}\frac{1}{\bar{z}^{n+1}}(z-1)^{n+1}, \tag{7.322}$$

where $\bar{z}$ denotes a value between z and 1. The latter series converges, as $n\to\infty$, if $z\in(0,2)$.

According to Corollary 7.4.1.3, for $j\in\mathbb{N}$ with $j\le n$, we are able to write

$$\begin{aligned}
&\left(\hat{\mathbb{E}}_x^{y,\epsilon}\left[e^{-(\epsilon/2)\int_0^1\beta(X_u)du}\right]-1\right)^j\\
&=\left(\sum_{k=1}^{n}\epsilon^k\phi_k(1,x,y)+\epsilon^n\mathcal{R}_\epsilon(x,y)\right)^j\\
&=\sum_{(z_1,\dots,z_n)\in S_{j,n}}\binom{j}{z_1,\dots,z_n}\epsilon^{z_1+\dots+nz_n}\phi_1(1,x,y)^{z_1}\cdot\ldots\cdot\phi_n(1,x,y)^{z_n}+\epsilon^n\mathcal{R}_\epsilon(x,y)\\
&=\sum_{k=j}^{n\cdot j}\epsilon^k\sum_{(z_1,\dots,z_n)\in M_{j,k,n}}\binom{j}{z_1,\dots,z_n}\phi_1(1,x,y)^{z_1}\cdot\ldots\cdot\phi_n(1,x,y)^{z_n}+\epsilon^n\mathcal{R}_\epsilon(x,y),
\end{aligned} \tag{7.323}$$

where in each line the expression $\mathcal{R}_\epsilon(x,y)$ stands for a different term that converges to zero as $\epsilon\to 0$. Due to the statements of Corollary 7.4.1.3, it is straightforward to show that, for each of these terms, convergence is uniform for (x,y) on compact subsets of $\mathbb{R}^2$.

Furthermore,

$$S_{j,n} = \left\{ (z_1, \ldots, z_n) \in \mathbb{N}_0^n \Big| \sum_{\nu=1}^{n} z_\nu = j \right\}, \tag{7.324}$$

that is $S_{j,n}$ is the set of all n-vectors of integers that sum up to j. Moreover, $M_{j,k,n}$ denotes the set

$$M_{j,k,n} = \left\{ (z_1, \ldots, z_n) \in \mathbb{N}_0^n \Big| \sum_{\nu=1}^{n} z_\nu = j,\ \sum_{\nu=1}^{n} \nu \cdot z_\nu = k \right\}. \tag{7.325}$$

Actually, as long as $k < n$, the set $M_{j,k,n}$ does not depend on n in the sense that $M_{j,k,n} = M_{j,k,k}$. This is easy to see. Let $(z_1, \ldots, z_n) \in \mathbb{N}_0^n$. If $z_\nu \neq 0$ for a $\nu > k$, then

$$z_1 + 2 \cdot z_2 + \ldots + n \cdot z_n > k. \tag{7.326}$$

If $n \geq k$, we set $M_{j,k,n} = M_{j,k,k} =: M_{j,k}$, and therefore we have

$$M_{j,k,n} = M_{j,k} = \left\{ (z_1, \ldots, z_k) \in \mathbb{N}_0^k \Big| \sum_{\nu=1}^{k} z_\nu = j,\ \sum_{\nu=1}^{k} \nu \cdot z_\nu = k \right\}. \tag{7.327}$$

Hence we are able to rewrite equation (7.323) in the following way

$$\begin{aligned}
&\left(\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2) \int_0^1 \beta(X_u) du} \right] - 1 \right)^j \\
&= \sum_{k=j}^{n} \epsilon^k \sum_{(z_1,\ldots,z_j) \in M_{j,k}} \binom{j}{z_1, \ldots, z_k} \phi_1(1,x,y)^{z_1} \cdot \ldots \cdot \phi_k(1,x,y)^{z_k} + \epsilon^n \mathcal{R}_\epsilon(x,y),
\end{aligned} \tag{7.328}$$

where $\mathcal{R}_\epsilon$ is (another) function that satisfies $\lim_{\epsilon \to 0} \mathcal{R}_\epsilon(x,y) = 0$, uniformly for (x,y) on compact subsets of $\mathbb{R}^2$. Combining (7.322) and (7.328), we find

$$\begin{aligned}
&\frac{1}{\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2) \int_0^1 \beta(X_u) du} \right]} \\
&= \sum_{j=0}^{n} (-1)^j \left(\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2) \int_0^1 \beta(X_u) du} \right] - 1 \right)^j \\
&\qquad + (-1)^{n+1} \frac{1}{\widetilde{\mathcal{X}}^{n+1}} \left(\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2) \int_0^1 \beta(X_u) du} \right] - 1 \right)^{n+1} \\
&= \sum_{j=0}^{n} (-1)^j \sum_{k=j}^{n} \epsilon^k \sum_{(z_1,\ldots,z_k) \in M_{j,k}} \binom{j}{z_1, \ldots, z_k} \phi_1(1,x,y)^{z_1} \cdot \ldots \cdot \phi_k(1,x,y)^{z_k} \\
&\qquad + \epsilon^n \mathcal{R}_\epsilon(x,y) + (-1)^{n+1} \frac{1}{\widetilde{\mathcal{X}}^{n+1}} \left(\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2) \int_0^1 \beta(X_u) du} \right] - 1 \right)^{n+1}
\end{aligned}$$

$$= \sum_{k=0}^{n} \epsilon^k \Phi^{(k)}(x,y) + \epsilon^n \mathcal{R}_\epsilon(x,y) + (-1)^{n+1} \frac{1}{\widetilde{\mathcal{X}}^{n+1}} \left(\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du} \right] - 1 \right)^{n+1}, \tag{7.329}$$

where again $\mathcal{R}_\epsilon(x,y)$ stands for a term that satisfies $\lim_{\epsilon\to 0} \mathcal{R}_\epsilon(x,y) = 0$. As before, convergence is uniform for (x,y) on compact subsets of $\mathbb{R}^2$. The term $\widetilde{\mathcal{X}}$ denotes a value between 1 and

$$\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du} \right]. \tag{7.330}$$

It remains to let ϵ tend to 0 and to estimate the last term on the right hand side of equation (7.329). Due to the definition of $\widetilde{\mathcal{X}}$ and since, on the one hand,

$$\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du} \right] - 1 = \epsilon \phi_1^\epsilon(1,x,y) + \epsilon^2 \mathcal{R}_\epsilon(x,y), \tag{7.331}$$

and, on the other hand,

$$\lim_{\epsilon\to 0} \hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du} \right] = 1, \tag{7.332}$$

uniformly for (x,y) on compact subsets of $\mathbb{R}^2$, this term satisfies

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^n} \left\{ (-1)^{n+1} \frac{1}{\widetilde{\mathcal{X}}^n} \left(\hat{\mathbb{E}}_x^{y,\epsilon} \left[e^{-(\epsilon/2)\int_0^1 \beta(X_u)du} \right] - 1 \right)^{n+1} \right\} = 0. \tag{7.333}$$

Again, convergence is uniform for (x,y) on compact subsets of $\mathbb{R}^2$, due to Corollary 7.4.1.3. Consequently, a combination of formulae (7.310), (7.320) and (7.329) yields the desired result. Altogether this proves part (i) of our proposition.

(ii) The second assertion is just a reformulation of the results we stated in Corollary 7.4.2.3 for $t = 1$. Again, we stress that under the measure $\hat{\mathbb{P}}_x^{y,\epsilon,t}$ the coordinate process X evolves like the Brownian bridge

$$x + \frac{u}{t}(y-x) + \sqrt{\epsilon}\left(B_u - \frac{u}{t} B_t \right), \quad u \in [0,t], \tag{7.334}$$

with B being standard Brownian motion of $\mathbb{R}$. □

In the following theorem, we summarize the results we have found .

Theorem 7.5.0.4. *Let X denote a diffusion that satisfies the following stochastic differential equation*

$$dX_u = \mu(X_u)du + \sigma(X_u)dB_u, \quad u \geq 0, \; X_0 = x. \tag{7.335}$$

Here, B denotes the standard Brownian motion of $\mathbb{R}$ and the coefficients μ and σ are assumed to belong to $\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ with σ uniformly bounded away from 0. Let $\bar{\mu}$ denote the *function*

$$z \longmapsto \bar{\mu}(z) = \left(\frac{\mu(F^{-1}(z))}{\sigma(F^{-1}(z))} - \frac{1}{2}\sigma'(F^{-1}(z)) \right), \quad z \in \mathbb{R}, \tag{7.336}$$

where F can be any primitive of $1/\sigma$, and assume that $\beta = \bar{\mu}' + \bar{\mu}^2$ satisfies Assumption 7.4.1.1 and Assumption 7.4.2.1. We denote with $\mathbb{P}_x$ the Markov measure on $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ making $\mathbb{P}_x[X_0 = x] = 1$. Let $x, h, y \in \mathbb{R}$, $x, y \leq h$ and let $n \in \mathbb{N}$. For $\epsilon > 0$, the probability of the first hitting time $\tau_h = \inf\{u > 0 \,|\, X_u \geq h\}$, conditional on $X_\epsilon = y$, satisfies the following expansion

$$\begin{aligned} &\mathbb{P}_x\big[\tau_h \leq \epsilon \,\big|\, X_\epsilon = y\big] \\ &= \exp\left(-\frac{2}{\epsilon} \int_x^h \frac{du}{\sigma(u)} \int_y^h \frac{du}{\sigma(u)} \right) \left\{ 1 + \sum_{k=1}^{n-1} \epsilon^k \Phi^{(k)}\Big(F(x), F(y)\Big) + \epsilon^n \Phi_\epsilon^{(n)}\Big(F(x), F(y)\Big) \right\} \\ &\qquad \times \left\{ 1 + \sum_{k=1}^{n-1} \epsilon^k \widetilde{\Phi}^{(k)}\Big(F(x), F(h), F(y)\Big) + \epsilon^n \widetilde{\Phi}_\epsilon^{(n)}\Big(F(x), F(h), F(y)\Big) \right\}. \end{aligned} \tag{7.337}$$

For $\xi, \eta \in \mathbb{R}$ and $k \in \mathbb{N}$, the functions $\Phi^{(k)}$ are defined via

$$\Phi^{(k)}(\xi, \eta) = \sum_{j=0}^{k} (-1)^j \sum_{(z_1, \ldots, z_k) \in M_{j,k}} \binom{j}{z_1, \ldots, z_k} \phi_1(1, \xi, \eta)^{z_1} \cdot \ldots \cdot \phi_k(1, \xi, \eta)^{z_k}, \tag{7.338}$$

where $M_{j,k}$ denotes the set (7.302). Here, $\phi_0 \equiv 1$ and, for $t \geq 0$, the functions ϕ_j are recursively defined by

$$\phi_j(t, \xi, \eta) = \int_0^t g_{j-1}\Big(t - u, \gamma_u^{(t,\xi,\eta)}\Big) du, \tag{7.339}$$

with

$$g_{j-1}(t, \xi, \eta) = \frac{1}{2} \frac{\partial^2}{\partial \xi^2} \phi_{j-1}(t, \xi, \eta) - \frac{1}{2} \beta(\xi) \phi_{j-1}(t, \xi, \eta), \tag{7.340}$$

and

$$\gamma_u^{(t,\xi,\eta)} = \xi + \frac{u}{t}(\eta - \xi), \quad u \in [0, t]. \tag{7.341}$$

Furthermore, for $\mathbb{h} \in \mathbb{R}$ with $\xi, \eta \leq \mathbb{h}$ and for $k \in \mathbb{N}$, $\widetilde{\Phi}^{(k)}(\xi, \mathbb{h}, \eta) = \tilde{\phi}_k(1, \xi, \mathbb{h}, \eta)$. Here, $\widetilde{\phi}_0 \equiv 1$ and, for $t \geq 0$, the functions $\tilde{\phi}_k$ are recursively defined by

$$\tilde{\phi}_k(t, \xi, \mathbb{h}, \eta) = \int_0^{t \frac{\mathbb{h} - \xi}{2\mathbb{h} - \xi - \eta}} \tilde{g}_{k-1}\Big(t - u, \rho_u^{(t,\xi,\mathbb{h},\eta)}, \mathbb{h}, \eta\Big) du + \phi_k\left(t \frac{\mathbb{h} - \eta}{2\mathbb{h} - \xi - \eta}, \mathbb{h}, \eta\right). \tag{7.342}$$

Again, ϕ_k is defined by (7.339), whereas $\tilde{g}_{k-1}$ is defined by

$$\tilde{g}_{k-1}(t,\xi,\mathbb{h},\eta) = \frac{1}{2}\frac{\partial^2}{\partial\xi^2}\tilde{\phi}_{k-1}(t,\xi,\mathbb{h},\eta) - \frac{1}{2}\beta(\xi)\tilde{\phi}_{k-1}(t,\xi,\mathbb{h},\eta), \tag{7.343}$$

and the path $\rho_\cdot^{(t,\xi,\mathbb{h},\eta)}$ is defined by

$$\rho_u^{(t,\xi,\mathbb{h},\eta)} = \begin{cases} \xi + \frac{u}{t}(2\mathbb{h} - \xi - \eta) & , \quad \text{if } \; 0 \le u \le t\frac{\mathbb{h}-\xi}{2\mathbb{h}-\xi-\eta}, \\ \eta + \frac{t-u}{t}(2\mathbb{h} - \xi - \eta) & , \quad \text{if } \; t\frac{\mathbb{h}-\xi}{2\mathbb{h}-\xi-\eta} < u \le t. \end{cases} \tag{7.344}$$

Finally, the functions $\Phi_\epsilon^{(n)}$ and $\widetilde{\Phi}_\epsilon^{(n)}$ on the right hand side of (7.337) satisfy

$$\lim_{\epsilon\to 0} \Phi_\epsilon^{(n)}(F(x), F(y)) = \Phi^{(n)}(F(x), F(y)), \tag{7.345}$$

uniformly for (x,y) on compact subsets of $\mathbb{R}^2$, and

$$\lim_{\epsilon\to 0} \widetilde{\Phi}_\epsilon^{(n)}(F(x), F(h), F(y)) = \widetilde{\Phi}^{(n)}(F(x), F(h), F(y)), \tag{7.346}$$

uniformly for (x,h,y) on compact subsets of $\{(x,h,y) \in \mathbb{R}^3 \,|\, x,y \le h\}$.

Proof. Let $Y_t = F(X_t)$, where F is a primitive of $1/\sigma$. By Itô's formula, the process Y satisfies the stochastic differential equation

$$dY_t = \left(\frac{\mu(F^{-1}(Y_t))}{\sigma(F^{-1}(Y_t))} - \frac{1}{2}\sigma'(F^{-1}(Y_t)) \right) dt + dB_t, \quad Y_0 = \xi = F(x), \; t \ge 0. \tag{7.347}$$

Before we proceed, let us emphasize that for the definition of the coefficients $\Phi^{(k)}$ and $\widetilde{\Phi}^{(k)}$ we considered the function $\beta = \bar{\mu}' + \bar{\mu}^2$, where $\bar{\mu} = (\mu/\sigma - \frac{1}{2}\sigma') \circ F^{-1}$ is the drift of the Lamperti transform (7.347).

The probability of the original process X, pinned by $X_\epsilon = y$, to cross level h equals the probability of Y, pinned by $\eta = F(X_\epsilon) = F(y)$, to cross level $\mathbb{h} = F(h)$. According to our previous notation, we consider the following two measures on $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R})$: first, let $\hat{\mathbb{Q}}_\xi^{\eta,\epsilon,F}$ denote the law of the Lamperti transform (7.347) starting at $Y_0 = \xi$ and pinned at $Y_\epsilon = \eta$. Furthermore, we denote with $\hat{\mathbb{P}}_\xi^{\eta,\epsilon}$ the law of the Brownian bridge

$$\xi + u(\eta - \xi) + \sqrt{\epsilon}(B_u - uB_1), \quad u \in [0,1], \tag{7.348}$$

where B is the standard Brownian motion of $\mathbb{R}$. This means that, under $\hat{\mathbb{Q}}_\xi^{\eta,\epsilon,F}$, the coordinate process X evolves like a solution to (7.347) and under $\hat{\mathbb{P}}_\xi^{\eta,\epsilon}$, X evolves like the Brownian bridge (7.348). By what we just stated and by (7.18), (7.300) and (7.304) we have

$$\mathbb{P}_x\left[\sup_{0\le s\le\epsilon} X_s \ge h \,\middle|\, X_\epsilon = y \right] = \hat{\mathbb{Q}}_\xi^{\eta,\epsilon,F}\left[\sup_{0\le s\le 1} X_s \ge \mathbb{h} \right]$$

$$= \hat{\mathbb{P}}_{\xi}^{\eta,\epsilon}\left[\sup_{0\le s\le 1} X_s \ge \mathbb{h}\right]\left\{1+\sum_{k=1}^{n-1}\epsilon^k\Phi^{(k)}(\xi,\eta)+\epsilon^n\Phi_\epsilon^{(n)}(\xi,\eta)\right\}$$
$$\times\left\{1+\sum_{k=1}^{n-1}\epsilon^k\widetilde{\Phi}^{(k)}(\xi,\mathbb{h},\eta)+\epsilon^n\widetilde{\Phi}_\epsilon^{(n)}(\xi,\mathbb{h},\eta)\right\}. \tag{7.349}$$

Note that we could have also written $\hat{\mathbb{Q}}_x^{y,\epsilon}\left[\sup_{0\le s\le 1} F(X_s)\ge \mathbb{h}\right]$ instead of $\hat{\mathbb{Q}}_\xi^{\eta,\epsilon,F}\left[\sup_{0\le s\le 1} X_s\ge \mathbb{h}\right]$ in the previous equation. Obviously,

$$\hat{\mathbb{P}}_{\xi}^{\eta,\epsilon}\left[\sup_{0\le s\le 1} X_s \ge \mathbb{h}\right] = \exp\left(-\frac{2}{\epsilon}(\mathbb{h}-\xi)(\mathbb{h}-\eta)\right). \tag{7.350}$$

The result now follows directly by Proposition 7.5.0.3. □

Another implication is stated in the following corollary. We replace the time variable ϵ by t, in order to achieve accordance with the notations of Chapter 2.

Corollary 7.5.0.5. *Let X be the process defined by* (7.335) *starting in x, and let the assumptions of Theorem 7.5.0.4 be satisfied. Moreover, we assume that the coefficients μ and σ are such that, for $t>0$, the ordinary transition probability density $p(t,x,y)$ of X exists. Let $h\in\mathbb{R}$ be fixed. Then the transition density $p^{(-\infty,h)}(t,x,y)$ of the diffusion X killed at h exists and it is given by*

$$p^{(-\infty,h)}(t,x,y) = p(t,x,y) - \mathbb{P}_x[\tau_h\le t\,|\,X_t=y]p(t,x,y). \tag{7.351}$$

The expression $\mathbb{P}_x[\tau_h\le t\,|\,X_t=y]$ can be expanded with respect to t according to formula (7.337)*, and the remainder terms in this expansion satisfy the asymptotics described at the end of Theorem 7.5.0.4.*

Proof. The fact that $p(t,x,y)$ exists implies the existence of $p^{(-\infty,h)}(t,x,y)$, where $p^{(-\infty,h)}(t,x,y)$ denotes the transition density of the diffusion X killed at h. For further details, see the outline concerning killed diffusions in Chapter 2. For a Borel set $A\in\mathcal{B}(\mathbb{R})$ with $\sup A\le h$, the following relation holds:

$$\int_A p^{(-\infty,h)}(t,x,y)dy = \mathbb{P}_x[H_t<h,\, X_t\in A] = \int_A\Big(1-\mathbb{P}_x[\tau_h\le t\,|\,X_t=y]\Big)p(t,x,y)dy. \tag{7.352}$$

Also see the discussion at the beginning of Chapter 3. Due to the results of Theorem 7.5.0.4, we are able to expand $\mathbb{P}_x[\tau_h\le t\,|\,X_t=y]$ with respect to t, and we are able to describe the asymptotical behavior of the remainder terms as $t\to 0$. □

We close this section with the following remark.

Remark 7.5.0.6. Note that the statement of Corollary 7.5.0.5 constitutes a large improvement compared to the representation for the transition density of a killed diffusion given by (2.18) in Chapter 2. Formula (7.351) can be approximated by means of the expansion described in Theorem 7.5.0.4, whereas – to our knowledge – there is no known result that allows to expand (2.18) directly. By a comparison of (2.18) and (7.351) we see that the following formula holds

$$\mathbb{P}_x\big[\tau_h \leq t \,|\, X_t = y\big] p(t,x,y) = \mathbb{E}_x\big[p(t-\tau_h, h, y)\, \mathbb{1}_{\{\tau_h < t\}}\big]. \tag{7.353}$$

8 Fourth Order Expansions

8.1 Introduction

In this chapter, the results of Chapter 7 are used to derive an expansion of $\mathbb{E}_x[g(H_t, X_t)]$ with respect to $\sqrt{t}$ including powers of four. Of course, there is no limitation in the sense that the techniques we apply allow for a calculation of higher order expansions as well. However, the calculations are quite tedious since every coefficient has to be calculated individually. But obviously, this is a significant improvement over the results of Section 5.2, where a standard approach yielded an expansion with respect to $\sqrt{t}$ with highest order $\sqrt{t}^2 = t$.

In order to state the main result of this section, an auxiliary result about the transition density of a diffusion is needed. On account of its importance, we dedicate the whole Section 8.2 to this issue. In Section 8.3 we are then going to calculate our expansion.

8.2 Series expansion of the transition density of a diffusion

The series expansion of transition densities we are going to present in this section was found by Aït-Sahalia, see [3] or [4]. Although there is an expansion in the multivariate case, we restrict ourselves to a description of the one-dimensional version.

We consider a diffusion X on the real line that satisfies the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x, \ t \geq 0. \tag{8.1}$$

Henceforth, we will assume that the diffusion coefficient σ constantly equals 1.

Assumption 8.2.0.7. *The diffusion coefficient satisfies $\sigma \equiv 1$.*

Note that this assumption is not a real restriction. For diffusion processes X with a more general diffusion coefficient the results stated below can be formulated for the Lamperti transform of X defined by

$$Y = F(X) = \int^X \frac{du}{\sigma(u)}. \tag{8.2}$$

Any primitive F of the function $\frac{1}{\sigma}$ may be selected, which means the constant of integration is irrelevant. If σ is bounded away from 0, the domain of Y is also $\mathbb{R}$. Recall that the process Y satisfies the stochastic differential equation

$$dY_t = \frac{\mu(F^{-1}(Y_t))}{\sigma(F^{-1}(Y_t))} - \frac{1}{2}\sigma'(F^{-1}(Y_t))dt + dB_t, \quad Y_0 = F(x),\ t \geq 0. \tag{8.3}$$

Let us go back to the core issue. In addition, we assume that the two following assumptions are satisfied.

Assumption 8.2.0.8. *The function $\mu(y)$ is differentiable in y infinitely many times.*

Assumption 8.2.0.9. *The function $\mu(y)$ and its derivatives with respect to y have at most polynomial growth near ∞. Moreover, we require $\lim_{y\to\infty \text{ or } -\infty} \lambda(y) < +\infty$, where $\lambda(y) = -\beta(y)/2 = -\left(\frac{\partial}{\partial y}\mu(y) + \mu(y)^2\right)/2$. Note that λ is not restricted from going to $-\infty$ near ∞. And finally, we assume that there exist constants $E > 0$ and $K > 0$, such that for all $y \leq -E$, $\mu(y) \geq Ky$, and for all $y \geq E$, $\mu(y) \leq Ky$.*

Let $p(t, x, y)$ denote the transition density of the process X starting in x. A Hermite-series approximation of p is given by

$$p^{(J)}(t, x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{J} \eta_j(t, x) H_j\left(\frac{x-y}{\sqrt{t}}\right), \tag{8.4}$$

where H_j is the j^{th} Hermite polynomial defined by

$$H_j(x) = \frac{(-1)^j}{j!} \exp\left(\frac{x^2}{2}\right) \frac{d^j}{dx^j} \exp\left(\frac{-x^2}{2}\right), \quad j \geq 1, \tag{8.5}$$

and the coefficients η_j satisfy

$$\eta_j(t, x) = \frac{\mathbb{E}_x\Big[H_j(t^{-1/2}(X_t - x))\Big]}{j!}, \quad j \geq 1. \tag{8.6}$$

For example, the first four Hermite-polynomials are

$$H_0(x) = 1,\ H_1(x) = x,\ H_2(x) = \frac{1}{2}(x^2 - 1),\ H_3(x) = \frac{1}{6}x^3 - \frac{1}{2}x. \tag{8.7}$$

Let $n \in \mathbb{N}$. In closed form the $(2n)^{th}$ Hermite polynomial is given by

$$H_{2n}(x) = \sum_{j=0}^{n} (-1)^{n-j} \frac{1}{(2j)!} \frac{1}{2^{n-j}(n-j)!} x^{2j}, \tag{8.8}$$

and in the odd case the $(2n+1)^{th}$ Hermite polynomial is given by

$$H_{2n+1}(x) = \sum_{j=0}^{n} \frac{1}{(2j+1)!} \frac{(-1)^{n-j}}{2^{n-j}(n-j)!} x^{2j+1}. \tag{8.9}$$

That means, H_j is a polynomial of order j. Its leading term is $\frac{x^j}{j!}$. Moreover, $H_{2n}(0) = \frac{(-1)^n}{2^n n!}$ in the even case and $H_{2n+1}(0) = 0$ in the odd case.

The coefficients η_j can be computed by making use of the following formula. Let $f : \mathbb{R} \to \mathbb{R}$ be a $(2n+1)$-times continuously differentiable function that does not grow too fast, then

$$\mathbb{E}_x[f(X_t)] = \sum_{j=0}^{n} \frac{t^n}{n!} \mathcal{A}^j f(a)|_{a=x} + o(t^n), \tag{8.10}$$

where $\mathcal{A}$ denotes the infinitesimal generator of the process X. Clearly, we have $\eta_0(t,x) = 1$. And the coefficients $\eta_1(t,x)$, $\eta_2(t,x)$ and $\eta_3(t,x)$ up to the order t^2 or $t^{5/2}$, respectively, are given by

$$\eta_1(t,x) = \mu(x)\sqrt{t} + \frac{1}{2}t^{3/2}\left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x)\right) + O(t^{5/2}), \tag{8.11}$$

$$\eta_2(t,x) = t(\mu(x)^2 + \mu'(x)) + O(t^2), \tag{8.12}$$

$$\eta_3(t,x) = t^{3/2}\left(\mu(x)^3 + \frac{3}{2}\mu(x)\mu'(x) + \frac{1}{4}\mu''(x)\right) + O(t^{5/2}). \tag{8.13}$$

The next theorem states that the approximation (8.4) converges uniformly to the transition density p of the process X.

Theorem 8.2.0.10. *On the Assumptions 8.2.0.7 - 8.2.0.9 there exists $\bar{t} > 0$, such that for every $t \in (0, \bar{t})$ and $(x,y) \in \mathbb{R}$*

$$p^J(t,x,y) \longrightarrow p(t,x,y), \tag{8.14}$$

as $J \to \infty$. In addition, the convergence is uniform for x on compact subsets of $\mathbb{R}$ and it is uniform for y on $\mathbb{R}$.

Proof. A concise proof can be found in [3]. □

Let us state an important proposition. It will turn out to be crucial in the sequel.

Proposition 8.2.0.11. *Let Assumptions 8.2.0.7 - 8.2.0.9 be satisfied. For $x \in \mathbb{R}$ fixed, the coefficients $\eta_j(t,x)$, $j \in \mathbb{N}$, defined by (8.6) satisfy*

$$\eta_j(t,x) = O(t^{j/2}). \tag{8.15}$$

Proof. The above assumptions ensure that for all $j \in \mathbb{N}$ the coefficients η_j exist. The result can be proved in a straightforward way, but this requires tedious calculations. We are going to give a proof that relies on Kolmogorov's equation.

First, note that

$$\eta_1(t,x) = \mathbb{E}_x\left[\frac{X_t - x}{\sqrt{t}}\Big|X_0 = x\right] = O(t^{1/2}), \tag{8.16}$$

and that each function $\eta_j(t,x)$ is continuous in t for fixed x. Moreover, it is continuously differentiable on each interval $(t_0, \bar{t})$ with $0 < t_0 < \bar{t}$. Consequently, we find that

$$\begin{aligned}
\frac{\partial}{\partial t}\eta_j(t,x) &= \frac{\partial}{\partial t}\mathbb{E}_x\left[H_j\left(\frac{X_t - x}{\sqrt{t}}\right)\right] \\
&= -\frac{1}{2}\int_{-\infty}^{\infty} H_j'\left(\frac{y-x}{\sqrt{t}}\right)\frac{y-x}{\sqrt{t}}\frac{1}{t}p(t,x,y)dy \\
&\qquad + \int_{-\infty}^{\infty} H_j\left(\frac{y-x}{\sqrt{t}}\right)\frac{\partial}{\partial t}p(t,x,y)dy \\
&= -\frac{1}{2}\int_{-\infty}^{\infty} H_{j-1}\left(\frac{y-x}{\sqrt{t}}\right)\frac{y-x}{\sqrt{t}}\frac{1}{t}p(t,x,y)dy \\
&\qquad + \int_{-\infty}^{\infty} \mathcal{A}H_j\left(\frac{\cdot - x}{\sqrt{t}}\right)(y)p(t,x,y)dy, \qquad (8.17)
\end{aligned}$$

where $\mathcal{A}$ denotes the infinitesimal generator of the diffusion X and provided that the interchange of differentiation and integration in the previous calculations is justified. We used the fact that $p(t,x,y)$ satisfies Kolmogorov's forward equation $\frac{\partial}{\partial t} = \mathcal{A}^*$ in the forward variable y. Since

$$H_{j-1}\left(\frac{y-x}{\sqrt{t}}\right)\frac{y-x}{\sqrt{t}} = jH_j\left(\frac{y-x}{\sqrt{t}}\right) + H_{j-2}\left(\frac{y-x}{\sqrt{t}}\right), \qquad (8.18)$$

and since

$$\begin{aligned}
\mathcal{A}H_j\left(\frac{\cdot - x}{\sqrt{t}}\right)(y) &= \mu(y)\frac{\partial}{\partial y}H_j\left(\frac{y-x}{\sqrt{t}}\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}H_j\left(\frac{y-x}{\sqrt{t}}\right) \\
&= \mu(y)H_{j-1}\left(\frac{y-x}{\sqrt{t}}\right)\frac{1}{\sqrt{t}} + \frac{1}{2}H_{j-2}\left(\frac{y-x}{\sqrt{t}}\right)\frac{1}{t}, \qquad (8.19)
\end{aligned}$$

we obtain the overall equation

$$\begin{aligned}
&\frac{\partial}{\partial t}\eta_j(t,x) \\
&= -\frac{1}{2}\int_{-\infty}^{\infty}\left\{jH_j\left(\frac{y-x}{\sqrt{t}}\right) + H_{j-2}\left(\frac{y-x}{\sqrt{t}}\right)\right\}\frac{1}{t}p(t,x,y)dy \\
&\quad + \int_{-\infty}^{\infty}\left\{\mu(y)H_{j-1}\left(\frac{y-x}{\sqrt{t}}\right)\frac{1}{\sqrt{t}} + \frac{1}{2}H_{j-2}\left(\frac{y-x}{\sqrt{t}}\right)\frac{1}{t}\right\}p(t,x,y)dy \\
&= -\frac{j}{2t}\int_{-\infty}^{\infty} H_j\left(\frac{y-x}{\sqrt{t}}\right)p(t,x,y)dy + \int_{-\infty}^{\infty}\mu(y)H_{j-1}\left(\frac{y-x}{\sqrt{t}}\right)\frac{1}{\sqrt{t}}p(t,x,y)dy \\
&= -\frac{j}{2t}\eta_j(t,x) + \int_{-\infty}^{\infty}\mu(y)H_{j-1}\left(\frac{y-x}{\sqrt{t}}\right)\frac{1}{\sqrt{t}}p(t,x,y)dy. \qquad (8.20)
\end{aligned}$$

If $\eta_{j-1}(t,x) \neq 0$, then there is a constant ζ, depending on μ only, such that

$$\frac{\partial}{\partial t}\eta_j(t,x) = -\frac{j}{2t}t\eta_j(t,x) + \zeta\frac{1}{\sqrt{t}}\eta_{j-1}(t,x). \qquad (8.21)$$

This can easily be shown by means of the mean value theorem. By the specific form of $\eta_{j-1}(t,x)$ in combination with the result of Theorem 8.2.0.10, we conclude that $t \mapsto \eta_{j-1}(t^2,x),\ t \in \mathbb{C}$, is a potential series with positive radius of convergence. Thus, it is a holomorphic function in a neighborhood of 0, and in this neighborhood there can be no accumulation point of zeros. Otherwise $\eta_{j-1}(t^2,x)$ would be identically equal to zero. If we assume that $\eta_{j-1}(t,x) = O(t^{(j-1)/2})$, then, by means of formula (8.21), we conclude that

$$n_j(t,x) = O(t^{j/2}). \tag{8.22}$$

The proof now follows by induction. □

8.3 Fourth order expansions of the joint moments

The aim of this section is to determine an expansion of

$$\mathbb{E}_x[g(H_t, X_t)], \tag{8.23}$$

with respect to $\sqrt{t}$ including powers of 4. Here, X denotes a diffusion defined by the stochastic differential equation (8.1) with constant diffusion coefficient $\sigma \equiv 1$ and H denotes its running maximum process defined by $H_t = \sup_{0 \le s \le t} X_s$. Furthermore, let $\tau_h = \{t > 0 \,|\, X_t \ge h\}$ be the first time the process X hits the level h. It is obvious to state that the joint distribution of (H_t, X_t) satisfies

$$\mathbb{P}_x[H_t < h,\ X_t \in A] = \int_A \mathbb{P}_x\left[\tau_h > t \,\middle|\, X_t = y\right] \mathbb{P}_x[X_t \in dy], \tag{8.24}$$

for any Borel set $A \in \mathcal{B}(\mathbb{R})$ with $\sup A \le h$. In Chapter 7 we found an expansion of $\mathbb{P}_x\left[\tau_h \le t \,\middle|\, X_t = y\right]$. An existence result for the joint density $f(t,x,h,y)$ of (H_t, X_t) is also available, see Section 3.2. Theoretically, this would enable us to calculate the joint density of (H_t, X_t), since we know expansions for the transition density $p(t,x,y)$. A possible way to expand p was depicted in the previous section. But, since we do not have regularity results or estimates for the joint density f, we consider a slight modification. Our proceeding in the present section is as follows: in the first Paragraph 8.3.1 we conduct some preliminary considerations. In the second Paragraph 8.3.2 some auxiliary technical tools are derived, before we state the main result in the last Paragraph 8.3.3.

8.3.1 Preliminary considerations

In order to derive an expansion of $\mathbb{E}_x[g(H_t, X_t)]$, it is reasonable to consider the Taylor expansion of the function g and then to calculate the moments $\mathbb{E}_x[(H_t - x)^m (Y_t - x)^n]$, $m, n \in \mathbb{N}_0$. For these moments a relatively simple representation exists. By formula (8.24), one easily obtains that, for fixed $y \in \mathbb{R}$, the expression

$$-\mathbb{P}_x\left[\tau_h \le t \,\middle|\, X_t = y\right] p(t,x,y) = -\mathbb{P}_x\left[H_t \ge h \,\middle|\, X_t = y\right] p(t,x,y) \tag{8.25}$$

is a primitive of $f(t,x,h,y)$ with respect to h. The function $f(t,x,h,y)$ denotes the joint density of the two-dimensional random variable (H_t, X_t), conditional on $X_0 = x$. For $m, n \in \mathbb{N}_0$, with $m \geq 1$, we have

$$\begin{aligned}
&\mathbb{E}_x[(H_t - x)^m (X_t - x)^n] \\
&\quad = \int_{-\infty}^{\infty} \mathbb{E}_x[(H_t - x)^m | X_t = y](y-x)^n p(t,x,y)dy \\
&\quad = \int_{-\infty}^{\infty} \int_{x \vee y}^{\infty} \mathbb{P}_x \left[\tau_h \leq t \,|\, X_t = y\right] m(h-x)^{m-1}\, dh\, (y-x)^n p(t,x,y)dy \\
&\qquad\qquad - \int_x^{\infty} (y-x)^{m+n} p(t,x,y)dy.
\end{aligned} \tag{8.26}$$

This formula can directly be verified by means of integration-by-parts. Note that the presence of the additional integral on the right hand side of the previous equation follows from the fact that

$$\mathbb{P}_x \left[\tau_{x \vee y} \leq t \,|\, X_t = y\right] = \mathbb{P}_x \left[H_t \geq x \vee y \,|\, X_t = y\right] \equiv 1. \tag{8.27}$$

The right hand side of (8.26) consists of terms that we are able to handle. The next Paragraph 8.3.2 is dedicated to the analysis of the first term on the right hand side of (8.26), which is

$$\int_{-\infty}^{\infty} \int_{x \vee y}^{\infty} \mathbb{P}_x \left[\tau_h \leq t \,|\, X_t = y\right] m(h-x)^{m-1}\, dh\, (y-x)^n p(t,x,y)dy. \tag{8.28}$$

We found an expansion of $\mathbb{P}_x \left[\tau_h \leq t \,|\, X_t = y\right]$ with respect to t in Chapter 7. Loosely speaking, this expansion was shown to converge uniformly for (t,x,h,y) on compact subsets of $[0,t_0] \times \{(x,h,y) \in \mathbb{R}^3 \,|\, x, y \leq h\}$, for sufficiently small $t_0 > 0$ and provided that Assumptions 7.4.1.1 and 7.4.2.1 are satisfied for the function $\beta = \mu' + \mu^2$. As usual, μ denotes the drift coefficient of the diffusion X. But since there is no result that states global convergence, we have to make an additional assumption.

Assumption 8.3.1.1. *Let us assume that the coefficient μ of the process X is such that there is an $h_0 > x$ so that, for all $h \geq h_0$, the quantity $\mathbb{P}_x[H_t \geq h]$ is exponentially negligible in the following sense: for all $j, k \in \mathbb{N}$,*

$$\lim_{t \to 0} \frac{\mathbb{E}_x\left[H_t^j \mathbb{1}_{\{H_t \geq h\}}\right]}{t^k} = \lim_{t \to 0} \frac{\int_h^\infty \mathbb{P}_x[H_t \geq a] j a^{j-1} da}{t^k} = 0. \tag{8.29}$$

Remark 8.3.1.2. Of course, in order to derive an expansion with respect to $\sqrt{t}$ including powers of $n \in \mathbb{N}$, it is sufficient to postulate that

$$\lim_{t \to 0} \frac{\int_h^\infty \mathbb{P}_x[H_t \geq a] j a^{j-1} da}{t^{n/2}} = 0, \tag{8.30}$$

for all $j \leq n$ and for all $h \geq h_0 > x$.

Remark 8.3.1.3. Clearly, Assumption 8.3.1.1 holds for Brownian motion. But let us derive some other examples. Suppose that the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x,\ t \geq 0, \tag{8.31}$$

has a weak solution X. If the coefficients $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ of X are uniformly bounded on $\mathbb{R}$ and if σ is uniformly elliptic, then X satisfies formula (8.29) in Assumption 8.3.1.1. This follows from the fact that in this special case Bernstein's inequality holds. Concretely, there exists a constant c such that

$$\mathbb{P}_x\left[\sup_{0\leq s\leq t} X_s \geq h\right] \leq c\exp\left(-\frac{(h-x)^2}{2ct}\right), \quad \forall t \geq 0. \tag{8.32}$$

This inequality is a direct consequence of Doob's inequality. For additional information, see e.g. Revuz and Yor [57].

Moreover, let us consider a process X that satisfies

$$dX_t = \mu(X_t)dt + dB_t, \quad X_0 = x,\ t \geq 0, \tag{8.33}$$

with a drift μ that is Lipschitz continuous and that satisfies the linear growth condition

$$|\mu(x)| \leq c\,(1+|x|)\,, \quad \forall x \in \mathbb{R}, \tag{8.34}$$

with a positive constant c. Without loss of generality we can assume that $c = 1$. Note that, for $\epsilon > 0$, the process $(U^\epsilon_t = X_{\epsilon t},\ t \geq 0)$ satisfies the stochastic differential equation

$$dU^\epsilon_t = \epsilon\mu(U^\epsilon_t)dt + \sqrt{\epsilon}dB_t, \quad U^\epsilon_0 = x,\ t \geq 0. \tag{8.35}$$

Trivially, for $0 < \epsilon \leq 1$,

$$x + \epsilon\int_0^t \mu(U^\epsilon_s)ds + \sqrt{\epsilon}B_t \ \leq\ x + \int_0^t (1+|U^\epsilon_s|)ds + \sqrt{\epsilon}B_t, \quad \forall t \geq 0. \tag{8.36}$$

We denote the process on the right hand side with $\bar{U}^\epsilon_t$, that is

$$\bar{U}^\epsilon_t = x + \int_0^t (1+|U^\epsilon_s|)ds + \sqrt{\epsilon}B_t, \quad t \geq 0. \tag{8.37}$$

We obtain the estimate

$$\left|\bar{U}^\epsilon_t\right| \leq |x| + \int_0^t (1+|U^\epsilon_s|)ds + \sqrt{\epsilon}|B_t|, \quad t \geq 0. \tag{8.38}$$

And by Gronwall's lemma we find the estimate

$$\sup_{0\leq t\leq 1}\left|\bar{U}^\epsilon_t\right| \leq \tilde{c}\left\{1 + |x| + \sqrt{\epsilon}\sup_{0\leq t\leq 1}|B_t|\right\}, \tag{8.39}$$

with a suitable constant $\tilde{c} > 1$. For a proof of Gronwall's lemma see, for example, the book of Dembo and Zeitouni [19]. By Lemma 5.2.1 in [19] and for h sufficiently large, it follows that

$$\mathbb{P}_x \left[\sup_{0 \le t \le 1} \left| \bar{U}_t^\epsilon \right| \ge h \right] \le 4 \exp\left(-\frac{(h/\tilde{c} - |x| - 1)^2}{2\epsilon} \right). \tag{8.40}$$

By our definitions, the following relation holds

$$\left\{ \sup_{0 \le t \le \epsilon} X_t \ge h \right\} = \left\{ \sup_{0 \le t \le 1} X_{t\epsilon} \ge h \right\} \subset \left\{ \sup_{0 \le t \le 1} |U_t^\epsilon| \ge h \right\}, \tag{8.41}$$

and hence, we conclude that a diffusion defined by (8.33), with a drift μ that satisfies (8.34), satisfies Assumption 8.3.1.1. Particularly, this includes the Ornstein-Uhlenbeck process which is defined by the stochastic differential equation

$$dX_t = \theta X_t dt + dB_t, \quad X_0 = x,\ t \ge 0, \tag{8.42}$$

where θ is a real valued parameter.

After what we have just stated, Assumption 8.3.1.1 is reasonable and it is satisfied for a certain class of processes. Particularly, if Assumptions 8.2.0.7 and 8.2.0.9 are satisfied, Assumption 8.3.1.1 is redundant.

In the next section, we will present some auxiliary results that are necessary to estimate the remainder term in the expansion of $\mathbb{E}_x[g(H_t, X_t)]$.

8.3.2 Auxiliary results

Throughout this section S denotes a compact subset of $\mathbb{R}^2$ that contains a sufficiently large neighborhood of (x, x). According to the expansion given in Theorem 7.5.0.4, we are able to write the restriction of the integral (8.28) to the set S in the following way

$$\begin{aligned}
&\int_{\mathbb{R}} \int_{x \vee y}^{\infty} \mathbb{1}_S \mathbb{P}_x \left[\tau_h \le t \,|\, X_t = y \right] m(h-x)^{m-1}\, dh\, (y-x)^n p(t,x,y) dy \\
&= \int_{-\infty}^{\infty} \int_{x \vee y}^{\infty} \mathbb{1}_S \exp\left(-2 \frac{(h-x)(h-y)}{t} \right) \left\{ 1 + t\Phi^{(1)}(x,y) + t^2 \Phi_t^{(2)}(x,y) \right\} \\
&\quad \times \left\{ 1 + t\widetilde{\Phi}^{(1)}(x,h,y) + t^2 \widetilde{\Phi}_t^{(2)}(x,h,y) \right\} m(h-x)^{m-1} dh (y-x)^n p(t,x,y) dy.
\end{aligned} \tag{8.43}$$

We split the previous integral into the two integrals

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{x \vee y}^{\infty} \mathbb{1}_S \exp\left(-2 \frac{(h-x)(h-y)}{t} \right) \left\{ 1 + t\Phi^{(1)}(x,y) + t\widetilde{\Phi}^{(1)}(x,h,y) \right\} \\
&\qquad \times\ m(h-x)^{m-1} dh (y-x)^n p(t,x,y) dy
\end{aligned} \tag{8.44}$$

and

$$
\begin{aligned}
t^2 \int_{-\infty}^{\infty} \int_{x\vee y}^{\infty} \mathbb{1}_S \exp\left(-2\frac{(h-x)(h-y)}{t}\right) \mathcal{R}^{(2)}(t,x,h,y) \\
\times\ m(h-x)^{m-1} dh (y-x)^n p(t,x,y) dy, \qquad (8.45)
\end{aligned}
$$

where the remainder term $\mathcal{R}^{(2)}$ is given by

$$
\begin{aligned}
\mathcal{R}^{(2)}(t,x,h,y) = \Big\{ \Phi_t^{(2)}(x,y)\left(1 + t\widetilde{\Phi}^{(1)}(x,h,y)\right) + \widetilde{\Phi}_t^{(2)}(x,h,y)\left(1 + t\Phi^{(1)}(x,y)\right) \\
+\ \left(\Phi_t^{(2)}(x,y) + \widetilde{\Phi}_t^{(2)}(x,h,y)\right)\Big\}. \qquad (8.46)
\end{aligned}
$$

Note that the functions $\Phi^{(1)}$, $\widetilde{\Phi}^{(1)}$, $\Phi^{(2)}$ and $\widetilde{\Phi}^{(2)}$ recursively depend on the potential $\beta = \mu' + \mu^2$. First, we are going to analyze the integral (8.45). We will show that (8.45) belongs to $O(t^{5/2})$ if $m \in \mathbb{N}$, $n \in \mathbb{N}_0$. In order to simplify our notations, for $\kappa, m, n \in \mathbb{N}_0$, we introduce the expression

$$
\begin{aligned}
\Xi(\kappa,m,n) = \frac{2}{t}\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \int_{x\vee y}^{\infty} \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) \\
\times\ (2h-y-x)^\kappa (h-x)^m (y-x)^n dh dy. \qquad (8.47)
\end{aligned}
$$

Recall the definition of the Hermite-polynomial expansion for $p(t,x,y)$ in (8.4). The first term in this expansion is the Gaussian density

$$
y \longmapsto \frac{1}{\sqrt{2\pi t}}\left(-\frac{(y-x)^2}{2t}\right). \qquad (8.48)
$$

Therefore, it is immediately clear that the term (8.47) plays a crucial role for our analysis. Fortunately it is easy to analyze. We obtain the following lemma.

Lemma 8.3.2.1. *Let $\kappa, m, n \in \mathbb{N}_0$ and let Ξ be the function defined in* (8.47). *Then*

$$
\Xi(\kappa,m,n) = O(t^{(\kappa+m+n)/2-1/2}). \qquad (8.49)
$$

Proof. It follows by direct calculations that for $\gamma \in \mathbb{N}_0$,

$$
\begin{aligned}
&\Xi(0,0,\gamma) \\
&= \frac{2}{t}\frac{1}{\sqrt{2\pi t}} \int_{x}^{\infty} \int_{-\infty}^{h} \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (h-x)^\gamma dy dh \\
&= \frac{2^{1+\frac{1}{2}(-1+\gamma)}\Gamma\left[1+\frac{\gamma}{2}\right]}{\sqrt{\pi}(1+\gamma)}\ t^{(\gamma-1)/2}, \qquad (8.50)
\end{aligned}
$$

where Γ denotes the Gamma-function

$$
\Gamma(x) = \int_0^\infty t^{x-1}\exp(-t)dt, \quad x > 0. \qquad (8.51)
$$

Since $2h - x - y = 2(h-x) - (y-x)$, we find

$$\Xi(\kappa, m, n) = \sum_{j=0}^{\kappa} 2\binom{\kappa}{j}(-1)^{\kappa-j}\,\Xi(0, m+j, n)\,\Xi(0, m, n+\kappa-j). \tag{8.52}$$

Finally, it follows by integration-by-parts that

$$\begin{aligned}
&\int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(h-x)^m (y-x)^n dydh \\
&= \frac{t}{2}\int_x^\infty \exp\left(-\frac{(h-x)^2}{2t}\right)(h-x)^{m+n-1} dh \\
&+ \frac{1}{4}\int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(h-x)^{m-1}(y-x)^{n+1} dydh \\
&- \frac{t}{2}n\int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(h-x)^{m-1}(y-x)^{n-1} dydh.
\end{aligned} \tag{8.53}$$

This yields the overall recursion

$$\begin{aligned}
\Xi(0, m, n) =& \frac{1}{\sqrt{2\pi t}}\int_x^\infty \exp\left(-\frac{(h-x)^2}{2t}\right)(h-x)^{m+n-1} dh \\
&+ \frac{1}{4}\Xi(0, m-1, n+1) - \beta\frac{t}{2}\Xi(0, m-1, n-1).
\end{aligned} \tag{8.54}$$

The first term on the right hand side of the latter equation belongs to $O(t^{(m+n)/2-1/2})$. With this and (8.50), the assertion follows easily by induction. □

We are now ready to state an important result which allows us to estimate the integral (8.45).

Proposition 8.3.2.2. *Let $x \in \mathbb{R}$ be fixed and let μ satisfy the Assumptions 8.2.0.7 - 8.2.0.9. If the function $\beta : \mathbb{R} \to \mathbb{R}$ defined by $\beta = \mu' + \mu^2$ satisfies the Assumptions 7.4.1.1 and 7.4.2.1, then, for a compact subset $S \subset \mathbb{R}^2$ that contains a neighborhood of (x, x), the expression* (8.45) *belongs to $O(t^{5/2})$ for $m, n \in \mathbb{N}_0$, $m \geq 1$.*

Proof. By means of Theorem 7.5.0.4, we infer that $\mathcal{R}^{(2)}(t, x, h, y)$ converges to 0 as $t \to 0$. Convergence is uniform for (h, y) on S. Consequently, for each $t_0 > 0$, there is a constant K_0, possibly depending on S, such that

$$\begin{aligned}
\Big|\int_{-\infty}^\infty \int_{x\vee y}^\infty \mathbb{1}_S \exp\left(-2\frac{(h-x)(h-y)}{t}\right)\mathcal{R}^{(2)}(t, x, h, y) \\
\times m(h-x)^{m-1} dh (y-x)^n p(t, x, y) dy\Big|
\end{aligned}$$

$$\leq K_0 \int_{-\infty}^{\infty} \int_{x\vee y}^{\infty} \mathbb{1}_S \exp\left(-2\frac{(h-x)(h-y)}{t}\right) m(h-x)^{m-1} dh |y-x|^n p(t,x,y) dy \tag{8.55}$$

for all $0 \leq t < t_0$. Our assumptions ensure that the expansion (8.4) converges to $p(t,x,y)$, uniformly for y on compact sets and if t is sufficiently small. Therefore, we are allowed to replace $p(t,x,y)$ in the previous inequality (8.55) by the normal density $1/\sqrt{2\pi t}\exp(-(y-x)^2/2t)$, times a constant factor, in order to obtain the overall estimate:

$$\begin{aligned}
&\Big| \int_{-\infty}^{\infty} \int_{x\vee y}^{\infty} \mathbb{1}_S \exp\left(-2\frac{(h-x)(h-y)}{t}\right) \mathcal{R}^{(2)}(t,x,h,y) \\
&\qquad\qquad\qquad \times m(h-x)^{m-1} dh (y-x)^n p(t,x,y) dy \Big| \\
&\quad \leq K_1 \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \int_{x\vee y}^{\infty} \mathbb{1}_S \exp\left(-2\frac{(h-x)(h-y)}{t} - \frac{(y-x)^2}{2t}\right) \\
&\qquad\qquad\qquad \times m(h-x)^{m-1} dh |y-x|^n dy.
\end{aligned} \tag{8.56}$$

Here, K_1 denotes another sufficiently large, positive constant, possibly depending on S, and the latter inequality (8.56) holds for all $0 \leq t < t_1$, with t_1 sufficiently small. From Lemma 8.3.2.1 we infer that the right hand side of (8.56) belongs to $O(t^{1/2})$. If n is even, this follows directly. If n is odd, apply Hölder's inequality. The assertion of our proposition follows, due to the extra factor t^2 in (8.45). □

From Proposition 8.3.2.2, we can infer the following result.

Corollary 8.3.2.3. *On the assumptions of Proposition 8.3.2.2 and for $m \in \mathbb{N}$, $n \in \mathbb{N}_0$, we have the following expansion*

$$\begin{aligned}
&\mathbb{E}_x[(H_t-x)^m (X_t-x)^n] \\
&\quad = \int_{-\infty}^{\infty} \int_{x\vee y}^{\infty} \exp\left(-2\frac{(h-x)(h-y)}{t}\right) \left\{1 + t\Phi^{(1)}(x,y) + t\widetilde{\Phi}^{(1)}(x,h,y)\right\} \\
&\qquad\qquad\qquad \times\ m(h-x)^{m-1} dh (y-x)^n p(t,x,y) dy \\
&\qquad - \int_{x\vee y}^{\infty} (y-x)^{m+n} p(t,x,y) dy +\ O(t^{5/2}).
\end{aligned} \tag{8.57}$$

Proof. First, let us state that for a compact set $S \subset \mathbb{R}^2$ that contains a neighborhood of (x,x), the expectation

$$\mathbb{E}_x\Big[(H_t-x)^m (X_t-x)^n \mathbb{1}_{\{(H_t,X_t)\in S^c\}}\Big] \tag{8.58}$$

is exponentially negligible, due to the Assumptions 8.2.0.7 and 8.2.0.9 – recall the discussion at the end of Remark 8.3.1.2. So is the term

$$\int_{x\vee y}^{\infty} \mathbb{1}_{S^c}(y-x)^{m+n} p(t,x,y)dy. \tag{8.59}$$

On the other hand

$$\begin{aligned}
&\mathbb{E}_x[(H_t-x)^m(X_t-x)^n \mathbb{1}_{\{(H_t,X_t)\in S\}}] \\
&\quad = \int_{-\infty}^{\infty}\int_{x\vee y}^{\infty} \exp\left(-2\frac{(h-x)(h-y)}{t}\right)\left\{1+t\Phi^{(1)}(x,y)+t\widetilde{\Phi}^{(1)}(x,h,y)\right\} \\
&\qquad\qquad\qquad\qquad \times\ m(h-x)^{m-1}dh(y-x)^n p(t,x,y)dy \\
&\qquad - \int_{-\infty}^{\infty}\int_{x\vee y}^{\infty} \mathbb{1}_{S^c}\exp\left(-2\frac{(h-x)(h-y)}{t}\right)\left\{1+t\Phi^{(1)}(x,y)+t\widetilde{\Phi}^{(1)}(x,h,y)\right\} \\
&\qquad\qquad\qquad\qquad \times\ m(h-x)^{m-1}dh(y-x)^n p(t,x,y)dy \\
&\qquad - \int_{x\vee y}^{\infty}(y-x)^{m+n}p(t,x,y)dy + \int_{x\vee y}^{\infty}\mathbb{1}_{S^c}(y-x)^{m+n}p(t,x,y)dy \\
&\qquad\quad +\ O(t^{5/2}).
\end{aligned} \tag{8.60}$$

The two integrals in the previous equation, that are restricted to the set S^c, are exponentially negligible and consequently the result follows. □

Let us now continue with the analysis of the remaining terms. Set $\Psi(x,h,y) = \Phi^{(1)}(x,y)+\widetilde{\Phi}^{(1)}(x,h,y)$. The function Ψ is given by

$$\Psi(x,h,y) = \begin{cases} \frac{1}{2}\left(\frac{\int_x^y \beta(u)du}{y-x} - \frac{\int_x^h \beta(u)du+\int_y^h \beta(u)du}{2h-x-y}\right) & \text{, if } x\neq y, \\ \frac{1}{2}\left(\beta(x) - \frac{\int_x^h \beta(u)du}{h-x}\right) & \text{, if } x=y, \end{cases} \tag{8.61}$$

where β in turn is given by $\beta = \mu^2+\mu'$. According to Corollary 8.3.2.3, it remains to calculate the integral

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{x\vee y}^{\infty}\exp\left(-2\frac{(h-x)(h-y)}{t}\right)\{1+t\Psi(x,h,y)\} \\
&\qquad\qquad\qquad \times\ m(h-x)^{m-1}dh(y-x)^n p(t,x,y)dy,
\end{aligned} \tag{8.62}$$

in order to find the fourth order expansion of $\mathbb{E}_x[(H_t-x)^m(X_t-x)^n]$ with respect to $\sqrt{t}$. We state an auxiliary Lemma that helps us to estimate the remainder terms of the integral (8.62).

Lemma 8.3.2.4. *Let $\beta = \mu'+\mu^2 : \mathbb{R}\to\mathbb{R}$ be three times continuously differentiable.*

Then the function Ψ *defined by* (8.61) *satisfies*

$$\Psi(x,h,y) = O((y-x)^2) + O((2h-x-y)^2) + O((h-x)^2), \tag{8.63}$$

where the notation $O\big((2h-x-y)^2\big)$ *means that*

$$\lim_{y\to x}\lim_{h\to x}\frac{O\big((2h-x-y)^2\big)}{(2h-x-y)^2} = \lim_{h\to x}\lim_{y\to x}\frac{O\big((2h-x-y)^2\big)}{(2h-x-y)^2} = const. \tag{8.64}$$

Proof. If β is three times continuously differentiable, a Taylor-expansion of the integrals in (8.61), first around y and h and then around x, yields

$$\begin{aligned}
&\Psi(x,h,y)\\
&= \frac{1}{2}\beta(y) - \frac{1}{4}(y-x)\beta'(y) + O((y-x)^2)\\
&\quad - \left[\frac{1}{2}\beta(h) - \frac{1}{4}(2h-x-y)\beta'(h) + O\big((2h-x-y)^2\big)\right]\\
&= \frac{1}{2}\beta(x) + \frac{1}{2}\beta'(x)(y-x) - \frac{1}{4}\beta'(x)(y-x) - \frac{1}{4}\beta''(x)(y-x)^2 + O((y-x)^2)\\
&\quad - \frac{1}{2}\left[\beta(x) + \beta'(x)(h-x) + \frac{1}{2}\beta''(x)(h-x)^2 - \frac{1}{2}(2(h-x)-(y-x))\beta'(x)\right.\\
&\quad \left. -\frac{1}{2}(2(h-x)-(y-x))\beta''(x)(h-x) + O((h-x)^2)\right]\\
&\quad + O((2h-x-y)^2).
\end{aligned} \tag{8.65}$$

Most of the terms vanish and by collecting the remaining ones, we infer that

$$\Psi(x,h,y) = O((y-x)^2) + O((2h-x-y)^2) + O((h-x)^2), \tag{8.66}$$

which is the desired result. □

By the previous lemma, we are able to derive the following result.

Corollary 8.3.2.5. *Let* Ψ *be the function* (8.61). *Moreover, assume that* $\beta = \mu' + \mu^2$ *is three times continuously differentiable and that* β *and its derivatives satisfy a polynomial growth condition near infinity. Then, for* $m \in \mathbb{N}$, $n \in \mathbb{N}_0$,

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{x\vee y}^{\infty} \exp\left(-2\frac{(h-x)(h-y)}{t}\right)\{1 + t\Psi(x,h,y)\}\\
&\qquad\qquad\qquad\qquad \times\, m(h-x)^{m-1}dh(y-x)^n p(t,x,y)dy\\
&= \int_{-\infty}^{\infty}\int_{x\vee y}^{\infty} \exp\left(-2\frac{(h-x)(h-y)}{t}\right) m(h-x)^{m-1}dh(y-x)^n p(t,x,y)dy\\
&\quad + O(t^{5/2}).
\end{aligned} \tag{8.67}$$

Proof. Lemma 8.3.2.4 states that

$$\Psi(x,h,y) = O((h-x)^2) + O((y-x)^2) + O((h-x)(y-x)). \tag{8.68}$$

Therefore, by Lemma 8.3.2.1, we find that

$$\begin{aligned} t\int_x^\infty \int_{-\infty}^h & \frac{1}{\sqrt{2\pi t}} m(h-x)^{m-1}(y-x)^n \\ & \times \Psi(x,h,y)\exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) dydh = O(t^{2+(m+n)/2}), \end{aligned} \tag{8.69}$$

and the result follows. □

In order to simplify the analysis, note that

$$\begin{aligned} & \int_{-\infty}^\infty \int_{x\vee y}^\infty \exp\left(-2\frac{(h-x)(h-y)}{t}\right) m(h-x)^{m-1} dh (y-x)^n p(t,x,y) dy \\ & = 2\int_{-\infty}^\infty \int_{x\vee y}^\infty \exp\left(-2\frac{(h-x)(h-y)}{t}\right) \\ & \qquad \times \frac{(2h-x-y)}{t}(h-x)^m dh (y-x)^n p(t,x,y) dy. \\ & + \int_{x\vee y}^\infty (y-x)^{m+n} p(t,x,y) dy. \end{aligned} \tag{8.70}$$

Equivalence follows easily by integration-by-parts. The two terms on the right hand side of the previous equation are more convenient for our purpose than the integral on the left hand side. We are now able to state the final result that will enable us to calculate the expansion of $\mathbb{E}_x[(H_t-x)^m(X_t-x)^n]$ with respect to $\sqrt{t}$ including powers of 4.

Proposition 8.3.2.6. *Let μ satisfy the Assumptions 8.2.0.7 - 8.2.0.9. Moreover, assume that $\beta = \mu' + \mu^2$ satisfies the Assumptions 7.4.1.1 and 7.4.2.1 of Chapter 7. Then, for $m,n \in \mathbb{N}$,*

$$\begin{aligned} \mathbb{E}_x[(H_t-x)^m(X_t-x)^n] & \\ = 2\int_{-\infty}^\infty \int_{x\vee y}^\infty & \exp\left(-2\frac{(h-x)(h-y)}{t}\right)\frac{(2h-x-y)}{t} \\ & \times (h-x)^m dh (y-x)^n p(t,x,y) dy \\ + \ O(t^{5/2}). & \end{aligned} \tag{8.71}$$

Proof. The proof follows directly from Corollary 8.3.2.3 in combination with the result of Corollary 8.3.2.5 and formula (8.70). □

From the latter proposition we immediately infer the following corollary.

Corollary 8.3.2.7. *Let the assumptions of Proposition 8.3.2.6 be satisfied. Let $m, n, J \in \mathbb{N}_0$ and let Ξ be the function* (8.47)*. Then*

$$\begin{aligned}&\mathbb{E}_x[(H_t - x)^m (X_t - x)^n] \\ &= \sum_{j=0}^{J} \sum_{k=1}^{j} \frac{c_{j,k}}{\sqrt{t}^k} \Xi(1, m, n+k) \eta_j(t, x) + O\big(\max\{t^{5/2},\, t^{(J+1+m+n)/2}\}\big). \end{aligned} \tag{8.72}$$

The coefficients $c_{j,k}$ are inferred from the j^{th} Hermite polynomial H_j in the following way

$$c_{j,k} = \frac{d^k}{dx^k} H_j(x)\Big|_{x=0}. \tag{8.73}$$

Proof. First, let us note that, for integers κ, m, n, α, the order (with respect to t) of the integral

$$\begin{aligned}&\frac{1}{\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h (2h - x - y)^\kappa (h - x)^m (y - x)^n \\ &\qquad \times \left(\frac{y - x}{\sqrt{t}}\right)^\alpha \exp\left(-\frac{2(h - y)(h - x)}{t} - \frac{(y - x)^2}{2t}\right) dy dh \end{aligned} \tag{8.74}$$

does not depend on α. This can be verified by direct calculations or it simply follows from Lemma 8.3.2.1. As a result, the following integrals with respect to the Hermite polynomials satisfy

$$\begin{aligned}&\int_x^\infty \int_{-\infty}^h \frac{2(2h - x - y)}{t\sqrt{2\pi t}} (h - x)^m (y - x)^n H_j\left(\frac{h - x}{\sqrt{t}}\right) \\ &\qquad \times \exp\left(-\frac{2(h - y)(h - x)}{t} - \frac{(y - x)^2}{2t}\right) dy dh = O(t^{(m+n)/2}),\end{aligned} \tag{8.75}$$

for all $j \in \mathbb{N}$. Recall that the coefficients $\eta_j(t, x)$ belong to $O(t^{j/2})$. By what we have just stated, the integral on the right hand side of equation (8.71) is equivalent to

$$\begin{aligned}&\int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h - y)(h - x)}{t} - \frac{(y - x)^2}{2t}\right) \frac{(2h - x - y)}{t\sqrt{2\pi t}} (h - x)^m (y - x)^n \\ &\qquad \times \sum_{j=0}^{J} \eta_j(t, x) H_j\left(\frac{h - x}{\sqrt{t}}\right) dy dh,\end{aligned} \tag{8.76}$$

plus a term that belongs to $O(\max\{t^{5/2}, t^{(J+1+m+n)/2}\})$. The assertion follows. □

8.3.3 The main result

Clearly, a function $g \in \mathcal{C}^{k+1}(\mathbb{R}^2, \mathbb{R})$ has the following Taylor expansion

$$\begin{aligned} g(h, y) &= \sum_{j=0}^{k} \frac{1}{j!} d^{(j)} g(x, x)(h - x, y - x)^j + o(\|(h - x, y - x)\|^k) \\ &= T_k\, g((x, x), (h, y)) + o(\|(h - x, y - x)\|^k), \end{aligned} \tag{8.77}$$

where the Taylor polynomial $T_k\, g((x, x), (h, y))$ consists of the terms

$$\begin{aligned} &d^{(j)} g(x, x)(h - x, y - x)^j \\ &= \sum_{i_1=1}^{2} \cdots \sum_{i_j=1}^{2} \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_k}} f(x_1, x_2)\big|_{x_1=x, x_2=x} \\ &\qquad\qquad \times (h - x, y - x)e_{i_1} \cdots (h - x, y - x)e_{i_j} \end{aligned} \tag{8.78}$$

with

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{8.79}$$

Particularly, for $k = 4$, one obtains the Taylor polynomial $(h, y) \mapsto T_4\, g((x, x), (h, y))$ which is displayed in formula (10.153) in the Appendix. By the two formulae (8.77) and (10.153) we find

$$\begin{aligned} \mathbb{E}_x[g(H_t, X_t)] &= \int_x^{\infty} \int_{-\infty}^{h} T_4 g((x, x), (h, y)) f(t, x, h, y) dy dh \\ &\quad + \int_x^{\infty} \int_{-\infty}^{h} r((h - x), (y - x)) f(t, x, h, y) dy dh, \end{aligned} \tag{8.80}$$

where the remainder term $r((h-x), (y-x))$ belongs to $o(\|(h-x, y-x)\|^4)$, and provided that the involved moments exist. Assumptions 8.2.0.7 and 8.2.0.9 ensure the existence of the integrals on the right hand side of the expression (8.80) if g and its derivatives satisfy a polynomial growth condition.

Now, let us assume that the coefficient μ satisfies the Assumptions 8.2.0.7 - 8.2.0.9 and that $\beta = \mu' + \mu^2$ satisfies Assumption 7.4.1.1 and Assumption 7.4.2.1 of Chapter 7. Then, by means of formula (8.80), we can calculate the expansion of $\mathbb{E}_x[g(H_t, X_t)]$ with respect to $\sqrt{t}$ including powers of 4 and we can estimate the remainder term. The necessary technical results were stated in Proposition 8.3.2.6 and Corollary 8.3.2.7 in the previous paragraph. It remains to calculate the coefficients explicitly. This will be done in Appendix 10.4.2 for each coefficient separately. We end this section by stating the overall result.

Theorem 8.3.3.1. *Let $g \in \mathcal{C}^5(\mathbb{R}^2, \mathbb{R})$ and assume that every partial derivative of g has* *polynomial growth near infinity. Let the process X satisfy the differential equation*

$$dX_t = \mu(X_t)dt + dB_t, \quad X_0 = x, \ t \geq 0, \tag{8.81}$$

where B denotes the standard Brownian motion of $\mathbb{R}$. Let the coefficient μ satisfy the Assumptions 8.2.0.7 - 8.2.0.9 and let the potential $\beta = \mu' + \mu^2$ satisfy Assumption 7.4.1.1 and 7.4.2.1 of Chapter 7. Let H_t denote the maximum of X at time t. Then the following expansion with respect to $\sqrt{t}$ holds:

$$\begin{aligned}
&\mathbb{E}_x[g(H_t, X_t)] \\
&= g(x,x) \\
&\quad + g_{(1,0)}(x,x)\left\{\sqrt{\frac{2}{\pi}}\sqrt{t} + \frac{1}{2}\mu(x)t + \frac{1}{4}t^2\left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x)\right) + \frac{1}{3}\frac{1}{\sqrt{2\pi}}t^{3/2}\left(\mu'(x) + \mu(x)^2\right)\right\} \\
&\quad + g_{(0,1)}(x,x)\left\{\mu(x)t + \frac{1}{2}t^2\left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x)\right)\right\} \\
&\quad + g_{(2,0)}(x,x)\frac{1}{2}\left\{t + \frac{4}{3}\sqrt{\frac{2}{\pi}}\mu(x)t^{3/2} + \frac{1}{2}t^2(\mu'(x) + \mu(x)^2)\right\} \\
&\quad + g_{(0,2)}(x,x)\frac{1}{2}\left\{t + t^2(\mu'(x) + \mu(x)^2)\right\} \\
&\quad + g_{(1,1)}(x,x)\left\{\frac{1}{2}t + \frac{4}{3}\sqrt{\frac{2}{\pi}}\mu(x)t^{3/2} + \frac{1}{2}t^2(\mu'(x) + \mu(x)^2)\right\} \\
&\quad + g_{(3,0)}(x,x)\frac{1}{6}\left\{2\sqrt{\frac{2}{\pi}}t^{3/2} + \frac{9}{4}\mu(x)t^2\right\} + g_{(0,3)}(x,x)\frac{1}{2}\mu(x)t^2 \\
&\quad + g_{(2,1)}(x,x)\frac{1}{2}\left\{\frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2} + 2\mu(x)t^2\right\} + g_{(1,2)}(x,x)\frac{1}{2}\left\{\frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2} + \frac{3}{2}\mu(x)t^2\right\} \\
&\quad + g_{(4,0)}(x,x)\frac{1}{8}t^2 + \frac{1}{8}g_{(0,4)}(x,x)t^2 + g_{(3,1)}(x,x)\frac{3}{8}t^2 + g_{(1,3)}(x,x)\frac{1}{4}t^2 + g_{(2,2)}(x,x)\frac{1}{2}t^2 + O(t^{5/2}).
\end{aligned} \tag{8.82}$$

Proof. By means of Taylor's formula and Corollary 8.3.2.7, we find that

$$\begin{aligned}
&\mathbb{E}_x[g(H_t, X_t)] \\
&= g(x,x) + \sum_{\substack{m,n=0 \\ m+n\geq 1}}^{4} d_{m,n}\, g_{(m,n)}(x,x) \sum_{j=0}^{4-(m+n)} \sum_{k=1}^{j} \frac{c_{j,k}}{\sqrt{t}^k}\, \Xi(1, m, n+k)\eta_j(t,x) + O(t^{5/2}),
\end{aligned} \tag{8.83}$$

where the factors $d_{m,n}$ correspond to the respective coefficients in the Taylor polynomial (10.153) in Appendix 10.4.2. The terms

$$\sum_{j=0}^{4-(m+n)} \sum_{k=1}^{j} \frac{c_{j,k}}{\sqrt{t}^k}\, \Xi(1, m, n+k)\eta_j(t,x), \quad m,n = 0,\ldots,4, \ m+n \geq 1, \tag{8.84}$$

are also calculated in Appendix 10.4.2. The result follows by inserting the corresponding values into equation (8.83). □

Remark 8.3.3.2. Note that previous results can be retrieved from Theorem 5.2.3.5. Indeed, a comparison of the expansion (5.84) and the expansion (8.82) shows that the coefficients belonging to $\sqrt{t}$ and to t coincide in both expansions. Moreover, if $(h, y) \mapsto g(h, y)$ does not depend on h, then (8.82) coincides with the results of Aït-Sahalia, see [3].

9 Conclusions

Our thesis was dedicated to the analysis of statistical properties of the triplet (H_t, L_t, X_t), for a one-dimesional diffusion process X defined by a time-homogeneous stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x,\ t \geq 0. \tag{9.1}$$

The random variables L_t and H_t denoted the maximum and the minimum of the process X at time t. Formally, they were defined by

$$H_t = \sup_{0\leq s\leq t} X_s \quad \text{and} \quad L_t = \inf_{0\leq s\leq t} X_s. \tag{9.2}$$

In Section 3.2, we used Malliavin calculus to prove an existence result for the joint density of (H_t, L_t, X_t). The existence result was sufficient to define generalized martingale estimating functions for the coefficients $\mu(\cdot;\theta)$ and $\sigma(\cdot;\theta)$ in a parameterized one-dimensional diffusion model

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = x,\ t \geq 0,\ \theta \in \Theta \subset \mathbb{R}, \tag{9.3}$$

based on the observations

$$(H_{i\Delta}, L_{i\Delta}, X_{i\Delta}) = \left(\sup_{(i-1)\Delta\leq s\leq i\Delta} X_s,\ \inf_{(i-1)\Delta\leq s\leq i\Delta} X_s,\ X_{i\Delta} \right), \tag{9.4}$$

on equally spaced observation intervals $((i-1)\Delta, i\Delta]$, $i = 1, \ldots, n$, and for a fixed sampling frequency Δ. In Chapter 4, we proved consistency and asymptotic normality of the resulting estimators for the parameter θ, as $n \to \infty$. But these are purely theoretical results. The density $f(\Delta, x, h, l, y)$ of the triplet $(H_\Delta, L_\Delta, X_\Delta)$, conditional on $X_0 = x$, is not known explicitly in general and it is extremely hard to determine numerically.

This is the reason why we strove to find alternative inference methods. A standard approach led to a second order expansion of the expression $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$ for diffusion processes with constant diffusion coefficient $\sigma > 0$. The results were displayed at the end of Chapter 5. This approach did not allow an expansion including higher powers of $\sqrt{t}$. However, in Chapter 6, we saw that the second order expansion was already sufficient to state small-Δ-optimality results in a parameterized diffusion model, where the diffusion coefficient has the multiplicative structure $\sigma(\cdot,\theta) = \theta \cdot \sigma(\cdot)$. We concentrated mainly on the analysis of different classes of martingale estimating functions that consist of either only linear terms or only quadratic terms. For a fixed sample size

n, our results yielded lower bounds for the variance that are uniform as $\Delta = t \to 0$. The theoretical results showed that one can benefit from incorporating the maximum H and the minimum L into the analysis if the aim is to estimate the diffusion coefficient. A comparison with an ordinary model based on equidistant observations $X_{i\Delta}$, $i = 1, \ldots, n$, showed that a generalized model which takes the data $(H_{i\Delta}, L_{i\Delta}, X_{i\Delta})$, $i = 1, \ldots, n$, into account is superior when it comes to estimating $\sigma(\cdot;\theta)$. For a fixed sample size n, the asymptotic lower bounds of the variance, when $\Delta \to 0$, are significantly lower in the generalized model and we were also able to construct estimating functions that attain these lower bounds. Let us give an example. Theoretically, a strictly range-based estimating function, constructed from $(H_{i\Delta}, L_{i\Delta})$, $i = 1, \ldots, n$, has an asymptotic lower bound for the variance that is about 82 % lower than the one for the ordinary estimating function constructed from the equidistant observations $X_{i\Delta}$, $i = 1, \ldots, n$. Compare the paragraph "Assessment of the results for range-based MEFs" in Section 6.3.3. We also conducted a simulation study for the Ornstein-Uhlenbeck process which corroborated our theoretical findings, see Section 6.4.

An aspect that has not been examined so far is the case of simultaneous asymptotics. Typical asymptotics for this scenario are described by

$$n \to \infty, \ \Delta_n \to 0, \ n\Delta_n \to \infty. \tag{9.5}$$

For the case of ordinary estimating functions, this problem has been extensively studied. Consistency and asymptotic normality for the resulting estimators were proved, see Florens-Zmirou [26] or alternatively Yoshida [72] for related results. Kessler [44] used higher order expansions of the moments of the transition distribution to obtain estimators that are approximately normal also when Δ_n goes relatively slowly to zero as n tends to infinity. Very recent results concerning this subject were presented by Sørensen, see [67]. However, for our model no results have been obtained so far. Understanding the simultaneous asymptotics of generalized martingale estimating functions is a problem that remains to be solved.

In Section 6.4, different simulations for an Ornstein-Uhlenbeck process showed that, for small values of $\Delta = t$, the proposed estimators of the diffusion coefficient $\sigma(\cdot;\theta) = \theta$, inferred from approximately optimal generalized martingale estimating functions, perform well. The variances of the generalized range-based estimators were significantly lower than the variances of the corresponding estimator constructed from equidistant observations $X_{i\Delta}$, $i = 1, \ldots, n$. The effect was particularly visible for small values of Δ. However, for relatively large Δ, our range-based estimators had a large bias that was due to the fact that we used first and second order approximations to the underlying moments of the Ornstein-Uhlenbeck process to calculate the estimating functions. Furthermore, as we mentioned in Remark 6.3.2.9 in Paragraph 6.3.2, an analysis of the small-Δ-behavior of estimating functions, that are more complex than strictly linear or strictly quadratic estimating functions, requires higher order expansions of the term $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$. With a second order expansion, we are not even able to deal with estimating functions that have both linear and quadratic terms. These two issues clearly indicate the need for an overall expansion of the moments of the triplet (H_t, L_t, X_t). In Chapter 7, we presented an approach that was based on partial differential equation techniques and that resulted in an overall expansion of the hitting time probability of a pinned diffusion

$$\mathbb{P}_x[\tau_h \leq t \,|\, X_t = y], \tag{9.6}$$

where $\tau_h = \inf\{t > 0 \,|\, X_t \geq h\}$. Note, that our results directly extend the work of Baldi and Caramellino, see [6], who analyzed the asymptotical properties of (9.6) as t tends to 0.

Our expansions of the hitting time probability (9.6) can be used to expand $\mathbb{E}_x[g(H_t, X_t)]$. Some of the higher order terms were explicitly calculated in Chapter 8. But our analysis required relatively strict regularity assumptions on the coefficients μ and σ of the process X. In a nutshell, we had to postulate that the two series

$$1 + \sum_{i=1}^{\infty} \epsilon^i \phi_i(t, x, y) \quad \text{and} \quad 1 + \sum_{i=1}^{\infty} \epsilon^i \tilde{\phi}_i(t, x, h, y), \tag{9.7}$$

which are basically the main components of the quantity (9.6), should be uniformly convergent for (t, x, h, y) on compact subsets of the space $[0, T] \times \mathbb{R}^3$. See Assumption 7.4.1.1 and Assumption 7.4.2.1. The functions ϕ_k and $\tilde{\phi}_k$ are recursively defined and they depend on the potential $\beta = \bar{\mu}' + \bar{\mu}^2$ and its derivatives. The coefficient $\bar{\mu} : \mathbb{R} \to \mathbb{R}$, in turn, denotes the drift of the Lamperti transform of X. We were able to show that, if β is a quadratic polynomial, uniform convergence holds on compact subsets of $[0, T] \times \mathbb{R}^3$. Thus, in the case of an Ornstein-Uhlenbeck process, or in the case of any other process whose potential β satisfies suitable quadratic growth conditions, uniform convergence for both series in (9.7) is guaranteed. Consequently, the class of processes for which an expansion of the hitting time probability can be calculated is not empty. But the exact classification – or at least a better classification – of the functions $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$ for which the two series (9.7) converge uniformly is a non-trivial problem. This issue demands more research.

The above mentioned expansion of the joint probability of (H_t, X_t) is certainly the outstanding result of the present thesis. It remains to exploit this result for statistical purposes. By means of an overall expansion of $\mathbb{E}_x[g(H_t, X_t)]$, classes of generalized martingale estimating functions, that are more complex than the strictly linear or the strictly quadratic models we considered in Chapter 6, can be analyzed. As we have already stated, an important example we deem useful is the class of estimating functions that contains both linear and quadratic terms in the expressions

$$\Big\{h - \mathbb{E}_{x,\theta}[H_\Delta]\Big\} \quad \text{and} \quad \Big\{y - \mathbb{E}_{x,\theta}[X_\Delta]\Big\}. \tag{9.8}$$

But various other classes are imaginable. Martingale estimating functions could also be obtained from eigenfunctions of the infinitesimal generator $\mathcal{A}_\theta^{(H,X)}$ of the two-dimensional Markov process (H, X) in the parameterized diffusion model. Let us briefly explain what we have in mind. First, recall that an expansion of $\mathbb{E}_{x,\theta}[g(H_\Delta, X_\Delta)]$ contains $\sqrt{\Delta}$-terms. Therefore, the domain $\mathrm{dom}(\mathcal{A}_\theta^{(H,X)})$ of the operator $\mathcal{A}_\theta^{(H,X)}$ must be a subset of $\{g \in \mathcal{C}^{1,0}(\mathbb{R}^2, \mathbb{R}) \,|\, g_{1,0} \equiv 0\}$. On regularity assumptions it might be possible to find functions $\varphi_j(\cdot;\theta) \in \mathrm{dom}(\mathcal{A}_\theta^{(H,X)})$ and real values $\lambda_j(\theta)$, $j = 1, \ldots, N$, such that

$$\mathcal{A}_\theta^{(H,X)}\varphi_j(\cdot;\theta) = -\lambda_j(\theta)\varphi_j(\cdot;\theta) \tag{9.9}$$

and

$$\mathbb{E}_{x,\theta}\left[\varphi_j(H_\Delta, X_\Delta;\theta)\right] = e^{-\lambda_j(\theta)\Delta}\varphi_j(x, x;\theta). \tag{9.10}$$

A class of martingale estimating functions could then be constructed from the expressions

$$g_{spec}(\Delta, x, h, l, y;\theta) = \sum_{j=1}^{N} a(\Delta, x;\theta)\kappa_j^{spec}(\Delta, x, h, l, y;\theta) \tag{9.11}$$

with

$$\kappa_j^{spec}(\Delta, x, h, l, y;\theta) = \varphi_j(h, y;\theta) - e^{-\lambda_j(\theta)\Delta}\varphi_j(x, x;\theta). \tag{9.12}$$

The small-Δ-behavior of such complex classes of generalized estimating functions ought to be examined and to be compared to ordinary martingale estimating functions. There is reason to believe that the findings of Chapter 6 can still be improved.

Another unsolved problem is to determine an overall expansion of $\mathbb{E}_x[g(H_t, L_t, X_t)]$ with respect to $\sqrt{t}$. This goal will be particularly difficult to achieve. Two barrier hitting times of diffusions are significantly more complex mathematical constructs than one barrier hitting times. Even in the case of a Brownian motion with drift, the joint distribution of (H_t, L_t, X_t) is extremely hard to determine. It is not clear if a PDE approach, like the one we used to solve the one barrier problem determining (9.6), will put us in a position to find an overall expansion of the two barrier hitting probability of a pinned diffusion. Maybe an elaborate combination of the upper barrier problem and the lower barrier problem can be used to solve the two barrier problem that describes the quantity

$$\mathbb{P}_x[\tau_{[l,h]} \leq t \,|\, X_t = y], \tag{9.13}$$

where $\tau_{[l,h]} = \inf\{t > 0 \,|\, X_t \notin (l, h)\}$. Either way, a solution to this issue has to be found. Once again, more detailed research is required.

10 Appendix - Missing Proofs

10.1 Proofs of Chapter 5

We give the remaining proofs of Lemma 5.2.3.3 and Theorem 5.2.3.5.

Proof of Lemma 5.2.3.3. We apply Itô's formula iteratively to $g(H_t, L_t, X_t)$. Tedious, but straightforward calculations show that

$$
\begin{aligned}
&\mathbb{E}_x[g(H_t, L_t, X_t)] \\
&= g(x,x,x) + \mathbb{E}_x \int_0^t g_{1,0,0}(H_s, L_s, X_s) dH_s + \mathbb{E}_x \int_0^t g_{0,1,0}(H_s, L_s, X_s) dL_s \\
&\quad + \mathbb{E}_x \int_0^t g_{0,0,1}(H_s, L_s, X_s) dX_s + \frac{1}{2}\mathbb{E}_x \int_0^t g_{0,0,2}(H_s, L_s, X_s) d\langle X \rangle_s \\
&= g(x,x,x) \\
&+ \ \mathbb{E}_x \int_0^t g_{1,0,0}(x,x,x) dH_s + \mathbb{E}_x \int_0^t \int_0^s g_{2,0,0}(H_v, L_v, X_v) dH_v dH_s \\
&\quad + \ \mathbb{E}_x \int_0^t \int_0^s g_{1,1,0}(H_v, L_v, X_v) dL_v dH_s + \mathbb{E}_x \int_0^t \int_0^s g_{1,0,1}(H_v, L_v, X_v) dX_v dH_s \\
&\quad + \ \frac{1}{2}\mathbb{E}_x \int_0^t \int_0^s g_{1,0,2}(H_v, L_v, X_v) d\langle X \rangle_v dH_s \\
&+ \ \mathbb{E}_x \int_0^t g_{0,1,0}(x,x,x) dL_s + \mathbb{E}_x \int_0^t \int_0^s g_{1,1,0}(H_v, L_v, X_v) dH_v dL_s \\
&\quad + \ \mathbb{E}_x \int_0^t \int_0^s g_{0,2,0}(H_v, L_v, X_v) dL_v dL_s + \mathbb{E}_x \int_0^t \int_0^s g_{0,1,1}(H_v, L_v, X_v) dX_v dL_s \\
&\quad + \ \frac{1}{2}\mathbb{E}_x \int_0^t \int_0^s g_{0,1,2}(H_v, L_v, X_v) d\langle X \rangle_v dL_s \\
&+ \ \mathbb{E}_x \int_0^t g_{0,0,1}(x,x,x)\mu(x) ds + \mathbb{E}_x \int_0^t \int_0^s g_{1,0,1}(H_v, L_v, X_v) dH_v \mu(X_s) ds \\
&\quad + \ \mathbb{E}_x \int_0^t \int_0^s g_{0,1,1}(H_v, L_v, X_v) dL_v \, \mu(X_s) ds \\
&\quad + \ \mathbb{E}_x \int_0^t \int_0^s g_{0,0,2}(H_v, L_v, X_v) dX_v \mu(X_s) ds \\
&\quad + \ \frac{1}{2}\mathbb{E}_x \int_0^t \int_0^s g_{0,0,3}(H_v, L_v, X_v) d\langle X \rangle_v dX_s \\
&\quad + \ \frac{1}{2}\mathbb{E}_x \int_0^t g_{0,0,2}(x,x,x)\sigma^2(X_s) ds \\
&\qquad + \frac{1}{2}\mathbb{E}_x \int_0^t \int_0^s g_{1,0,2}(H_v, L_v, X_v) dH_v \, \sigma^2(X_s) ds
\end{aligned}
$$

$$+\frac{1}{2}\mathbb{E}_x\int_0^t\int_0^s g_{0,1,2}(H_v,L_v,X_v)dL_v\,\sigma^2(X_s)ds$$
$$+\frac{1}{2}\mathbb{E}_x\int_0^t\int_0^s g_{0,0,3}(H_v,L_v,X_v)dX_v\,\sigma^2(X_s)ds$$
$$+\frac{1}{4}\mathbb{E}_x\int_0^t\int_0^s g_{0,0,4}(H_v,L_v,X_v)d\langle X\rangle_v\,\sigma^2(X_s)ds. \tag{10.1}$$

In formula (10.1) we already made use of the fact that $d\langle X\rangle_t = \sigma^2(X_t)dt$. As a first step, we filter out the terms that appear on the right hand side of formula (5.82). It remains to show that each of the remaining terms in (10.1) belongs to $O(t^{3/2})$. We are going to consider these terms individually. Let us start with the term

$$\frac{1}{2}\mathbb{E}_x\int_0^t g_{0,0,2}(x,x,x)\sigma^2(X_s)ds = \frac{1}{2}\mathbb{E}_x\int_0^t g_{0,0,2}(x,x,x)\sigma^2(x)ds + \frac{1}{2}\mathbb{E}_x\int_0^t g_{0,0,2}(x,x,x)\left\{\sigma^2(X_s)-\sigma^2(x)\right\}ds. \tag{10.2}$$

By the fact that

$$\left|\sigma^2(X_s)-\sigma^2(x)\right| \le 2\left|\sigma(X_s)\right|\left|\sigma(X_s)-\sigma(x)\right| + 2\left|\sigma(X_s)-\sigma(x)\right|^2, \tag{10.3}$$

and by the Lipschitz continuity and the linear growth condition for σ, the remainder term on the right hand side of (10.2) is trivial to estimate. Now, let us consider the three terms

$$\mathbb{E}_x\int_0^t\int_0^s g_{1,0,2}(H_v,L_v,X_v)d\langle X\rangle_v dH_s, \tag{10.4}$$

$$\mathbb{E}_x\int_0^t\int_0^s g_{0,1,2}(H_v,L_v,X_v)d\langle X\rangle_v dL_s \tag{10.5}$$

and

$$\mathbb{E}_x\int_0^t\int_0^s g_{0,0,3}(H_v,L_v,X_v)d\langle X\rangle_v dX_s. \tag{10.6}$$

Let $\alpha,\beta,\gamma\in\mathbb{N}_0$. Then, by Cauchy-Schwarz' inequality and by the linear growth condition imposed upon σ, we find the estimate

$$\begin{aligned}
&\left|\mathbb{E}_x\int_0^t\int_0^s H_v^\alpha\, L_v^\beta\, X_v^\gamma d\langle X\rangle_v dH_s\right| \\
&\preceq t\,\mathbb{E}_x\left[H_t^{2+\alpha+\gamma}(2x-L_t)^\beta\int_0^t dH_s\right] \\
&\le t\sqrt{\mathbb{E}_x\Big[H_t^{4+2\alpha+2\gamma}(2x-L_t)^{2\beta}\Big]}\sqrt{\mathbb{E}_x\left[\left(\int_0^t dH_s\right)^2\right]}.
\end{aligned} \tag{10.7}$$

Since $\mathbb{E}_x\left[\left(\int_0^t dH_s\right)^2\right] = O(t)$ and since g and its partial derivatives have polynomial growth, it follows that

$$\mathbb{E}_x \int_0^t \int_0^s g_{1,0,2}(H_v, L_v, X_v) d\langle X\rangle_v dH_s = O(t^{3/2}). \tag{10.8}$$

The terms (10.5) and (10.6) can be bounded in an analogous way and thus, the three terms (10.4), (10.5) and (10.6) belong to $O(t^{3/2})$.

Next, we estimate the integrals of the form

$$\mathbb{E}_x \left[\int_0^t \int_0^s \gamma(H_u, L_u, X_u) dA_u \mu(X_s) ds\right] \tag{10.9}$$

or

$$\mathbb{E}_x \left[\int_0^t \int_0^s \gamma(H_u, L_u, X_u) dA_u \sigma^2(X_s) ds\right], \tag{10.10}$$

where $A \in \{H, L, X\}$ and where $\gamma : \mathbb{R} \to \mathbb{R}$ is a partial derivative of g. We can make use of similar arguments as above. Let $\alpha, \beta, \gamma \in \mathbb{N}_0$, then

$$\begin{aligned}\left|\mathbb{E}_x \int_0^t \int_0^s H_u^\alpha L_u^\beta X_u^\gamma \, dH_u ds\right| &\leq t \, \mathbb{E}_x \left[H_t^{2+\alpha+\gamma}(2x - L_u)^\beta \int_0^t dH_s\right] \\ &\leq t \sqrt{\mathbb{E}_x\left[H_t^{4+2\alpha+2\gamma}(2x - L_u)^{2\beta}\right]} \sqrt{\mathbb{E}_x \left[\left(\int_0^t dH_s\right)^2\right]}. \end{aligned} \tag{10.11}$$

The integrals

$$\mathbb{E}_x \int_0^t \int_0^s H_u^\alpha L_u^\beta X_u^\gamma \, dL_u ds, \tag{10.12}$$

and

$$\mathbb{E}_x \int_0^t \int_0^s H_u^\alpha L_u^\beta X_u^\gamma \, dX_u ds, \tag{10.13}$$

are estimated in a very similar way. Recall that μ and σ were assumed to satisfy a linear growth condition. Since g is of polynomial growth, this, together with the fact that $\mathbb{E}_x\left[\left(\int_0^t dH_s\right)^2\right] = O(t)$, shows that expressions of the form (10.9) or (10.10), respectively, belong to the class $O(t^{3/2})$.

Normally, we would have to further expand the remaining terms in formula (10.1). We are going to omit the concise outline, since another application of Itô's formula would make formula (10.1) even more confusing. Instead, we content ourselves with stating that, after another application of Itô's formula to the remaining terms on the right hand side of formula (10.1), one obtains terms of the following form

$$\mathbb{E}_x \int_0^t \int_0^s \int_0^v \gamma(H_u, L_u, X_u) dA_u^{(1)} dA_v^{(2)} dA_s^{(3)}, \tag{10.14}$$

where γ is a suitable derivative of g and $A^{(i)} \in \{H, L, X\}$ for $i = 1, 2, 3$. Again, let $\alpha, \beta, \gamma \in \mathbb{N}_0$ and note that

$$\left|\mathbb{E}_x \int_0^t \int_0^s \int_0^v H_u^\alpha L_u^\beta X_u^\gamma \, dH_u dH_v dH_s\right| \leq \mathbb{E}_x \left[H_t^{\alpha+\gamma}(2x - L_t)^\beta \int_0^t \int_0^s \int_0^v 1 dH_u dH_v dH_s\right]$$
$$\leq \sqrt{\mathbb{E}_x \left[H_t^{2\alpha+2\gamma}(2x - L_t)^{2\beta}\right]} \sqrt{\mathbb{E}_x \left[\left(\int_0^t \int_0^s \int_0^v 1 dH_u dH_v dH_s\right)^2\right]}. \tag{10.15}$$

Since $\mathbb{E}_x \left[\left(\int_0^t \int_0^s \int_0^v 1 dH_u dH_v dH_s\right)^2\right] = O(t^3)$, the expression

$$\mathbb{E}_x \int_0^t \int_0^s \int_0^v H_u^\alpha L_u^\beta X_u^\gamma \, dH_u dH_v dH_s \tag{10.16}$$

belongs to the class $O(t^{3/2})$. Terms of the form

$$\mathbb{E}_x \int_0^t \int_0^s \int_0^v H_u^\alpha L_u^\beta X_u^\gamma \, dA_u^{(1)} dA_v^{(2)} dA_s^{(3)}, \tag{10.17}$$

with different combinations of $A^{(i)} \in \{H, L, X\}$ are treated very similarly. Therefore, and since g has polynomial growth, expressions of the form (10.14) belong to $O(t^{3/2})$. In the previous step, we neglected several terms comprising the quadratic variation of X. But these terms are even easier to handle.

Finally, we have to estimate the term

$$\frac{1}{4}\mathbb{E}_x \int_0^t \int_0^s g_{0,0,4}(H_v, L_v, X_v) d\langle X\rangle_v \, \sigma^2(X_s) ds. \tag{10.18}$$

It is trivial to see that it must belong to $O(t^{3/2})$. As we already stated in Remark 5.2.3.4, a slight modification of the formula (10.1) shows that the fourth derivative of g with respect to x can be avoided. Indeed, the assumption $g \in \mathcal{C}^{3,3,3}(\mathbb{R}^3, \mathbb{R})$ is sufficient to state the result. This only requires a clever application of Taylor's formula. And since (10.18) is the only term that depends on a fourth order partial derivative of g, the analysis of the other terms in (10.1) basically remains unaffected.

Overall, we have shown that the remainder terms in formula (10.1) belong to the class $O(t^{3/2})$, which completes the proof of our lemma. □

Proof of Theorem 5.2.3.5. Due to the result of Lemma 5.2.3.3, it remains to calculate the expectations on the right hand side of formula (5.82). According to our previous notations, we denote with $\widetilde{X}_t$ the solution to

$$d\widetilde{X}_t = \mu(x)dt + \sigma dB_t, \quad \widetilde{X}_0 = x, \ t \geq 0. \tag{10.19}$$

Furthermore, let $\widetilde{H}_t$ and $\widetilde{L}_t$ denote the processes

$$\widetilde{H}_t = \sup_{0\le s\le t} \widetilde{X}_s, \quad \text{and} \quad \widetilde{L}_t = \inf_{0\le s\le t} \widetilde{X}_s, \tag{10.20}$$

respectively. We begin with the analysis of the cross-terms. Note, that

$$\mathbb{E}_x[H_t X_t] = \mathbb{E}_x[\widetilde{H}_t \widetilde{X}_t] + O(t^{3/2}). \tag{10.21}$$

This is a direct consequence of Lemma 5.2.1.4 and Theorem 5.2.1.6, since

$$\mathbb{E}_x\big[(H_t - x)(X_t - x)\big] = x^2 - x\mathbb{E}_x[H_t] - x\mathbb{E}_x[X_t] + \mathbb{E}_x[H_t X_t]. \tag{10.22}$$

By Itô's formula and by the estimate (10.21) we infer that

$$\mathbb{E}_x \int_0^t \int_0^s dX_v dH_s + \mathbb{E}_x \int_0^t \int_0^s dH_v dX_s = \mathbb{E}_0[\widetilde{H}_t \widetilde{X}_t] + O(t^{3/2}). \tag{10.23}$$

Clearly, $\mathbb{E}_0\left[\sup_{0\le s\le 1} B_t \cdot B_t\right] = \frac{1}{2}t$. By Corollary 5.2.1.4 it follows that

$$\mathbb{E}_0[\widetilde{H}_t \widetilde{X}_t] = \frac{1}{2}\sigma(x)^2 t + O(t^{3/2}), \tag{10.24}$$

which implies

$$\mathbb{E}_x \int_0^t \int_0^s dX_v dH_s + \mathbb{E}_x \int_0^t \int_0^s dH_v dX_s = \frac{1}{2}\sigma(x)^2 t + O(t^{3/2}). \tag{10.25}$$

Let us state that

$$\begin{aligned}
&\mathbb{E}_0\left[\left(\inf_{0\le s\le t} B_s\right) B_t\right] \\
&= \mathbb{E}_0\left[\left(\inf_{0\le s\le t}(-B_s)\right)(-B_t)\right] = \mathbb{E}_0\left[\left(-\inf_{0\le s\le t}(-B_s)\right) B_t\right] = \mathbb{E}_0\left[\left(\sup_{0\le s\le t} B_s\right) B_t\right].
\end{aligned} \tag{10.26}$$

Consequently, we can easily infer that likewise

$$\mathbb{E}_x \int_0^t \int_0^s dX_v dL_s + \mathbb{E}_x \int_0^t \int_0^s dL_v dX_s = \frac{1}{2}\sigma(x)^2 t + O(t^{3/2}). \tag{10.27}$$

And by arguments similar to the ones above, we find

$$\mathbb{E}_x \int_0^t \int_0^s dL_v dH_s + \mathbb{E}_x \int_0^t \int_0^s dH_v dL_s = (1 - 2\log 2)\sigma(x)^2 t + O(t^{3/2}). \tag{10.28}$$

Note that the formula $\mathbb{E}_0\left[\sup_{0\le s\le t} B_s \cdot \inf_{0\le s\le t} B_s\right] = (1-2\log 2)t$ was proved by Rogers and Shepp, see [62].

By Itô's formula and by the same arguments as above, one obtains that the moments of the process H must satisfy

$$\mathbb{E}_x\left[\int_0^t dH_s\right] = \mathbb{E}_0[\widetilde{H}_t] + O(t^{3/2}) = \sqrt{\frac{2}{\pi}}\sigma\sqrt{t} + \frac{1}{2}t\mu(x) + O(t^{3/2}) \tag{10.29}$$

and

$$\mathbb{E}_x\left[\int_0^t\int_0^s dH_v dH_s\right] = \frac{1}{2}\mathbb{E}_0[\widetilde{H}_t^2] + O(t^{3/2}) = \frac{1}{2}\sigma^2 t + O(t^{3/2}). \tag{10.30}$$

Finally, let us analyze the behaviour of the minimum process. The following property is pretty obvious

$$\inf_{0\le s\le t}\widetilde{X}_s = -\sup_{0\le s\le t}\{-\widetilde{X}_s\}. \tag{10.31}$$

On the other hand, for $x \in \mathbb{R}$ fixed, the processes

$$-\widetilde{X}_t = -\mu(x)t - \sigma B_t \tag{10.32}$$

and

$$-\mu(x)t + \sigma B_t \tag{10.33}$$

are equal in law. By Itô's formula, we find the expansions

$$\mathbb{E}_x\left[\int_0^t dL_s\right] = \mathbb{E}_0[\widetilde{L}_t] + O(t^{3/2}) = -\sqrt{\frac{2}{\pi}}\sigma\sqrt{t} + \frac{1}{2}\mu(x)t + O(t^{3/2}) \tag{10.34}$$

and

$$\mathbb{E}_x\left[\int_0^t\int_0^s dL_v dL_s\right] = \frac{1}{2}\mathbb{E}_0[\widetilde{L}_t^2] + O(t^{3/2}) = \frac{1}{2}\sigma^2 t + O(t^{3/2}). \tag{10.35}$$

By inserting the expressions we found into formula (5.82) in Lemma 5.2.3.3, we are able to infer the assertion. This completes the proof of Theorem 5.2.3.5. □

10.2 Proofs of Chapter 6

We prove Proposition 6.3.2.5 and Proposition 6.3.2.6.

Proof of Proposition 6.3.2.5. Because of (6.54), we have

$$\frac{\partial}{\partial h} g_{0,\theta}^2(x,h,x,x)\Big|_{h=x} = 2g_{0,\theta}(x,x,x,x)\frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x} = 0. \tag{10.36}$$

And similarly, we have

$$\frac{\partial}{\partial l} g_{0,\theta}^2(x,x,l,x)\Big|_{l=x} = \frac{\partial}{\partial y} g_{0,\theta}^2(x,x,x,y)\Big|_{y=x} = 0. \tag{10.37}$$

It remains to show that

$$\widetilde{g_{0,\theta}^2}(x,x,x,x) = 0. \tag{10.38}$$

But this is evident, since

$$\frac{\partial}{\partial s} g_{s,\theta}^2(x,x,x,x) = 2g_{s,\theta}(x,x,x,x)\frac{\partial}{\partial s} g_{s,\theta}(x,x,x,x), \tag{10.39}$$

which implies

$$\widetilde{g_{0,\theta}^2}(x,x,x,x) = 2g_{0,\theta}(x,x,x,x)\tilde{g}_{0,\theta}(x,x,x,x) = 0. \tag{10.40}$$

Therefore, the $\sqrt{\Delta}$-term in the expansion of $\mathbb{E}_{x,\theta}[g_{\Delta,\theta}^2]$ vanishes and the Δ-term becomes

$$\begin{aligned}
& -\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 g_{0,\theta}^2(x,x,x,x) + \mathcal{A}_\theta^{(1)} g_{0,\theta}^2(x,x,x,x) + \widetilde{\widetilde{g_{0,\theta}^2}}(x,x,x,x) \\
= \; & -\sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial h^2} g_{0,\theta}^2(x,h,x,x)\Big|_{h=x} \\
& - \sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial l^2} g_{0,\theta}^2(x,x,l,x)\Big|_{l=x} \\
& + 2\sigma(x;\theta)^2 \frac{2}{\pi}\frac{\partial^2}{\partial h \partial l} g_{0,\theta}^2(x,h,l,x)\Big|_{h=x,l=x} \\
& + \frac{1}{2}\sigma(x;\theta)^2 \frac{\partial^2}{\partial h^2} g_{0,\theta}^2(x,h,x,x)\Big|_{h=x} \\
& + \frac{1}{2}\sigma(x;\theta)^2 \frac{\partial^2}{\partial l^2} g_{0,\theta}^2(x,x,l,x)\Big|_{l=x} \\
& + (1-2\log 2)\sigma(x;\theta)^2 \frac{\partial^2}{\partial h \partial l} g_{0,\theta}^2(x,h,l,x)\Big|_{h=x,l=x} \\
& + \frac{1}{2}\sigma(x;\theta)^2 \frac{\partial^2}{\partial h \partial y} g_{0,\theta}^2(x,h,x,y)\Big|_{h=x,y=x} \\
& + \frac{1}{2}\sigma(x;\theta)^2 \frac{\partial^2}{\partial l \partial y} g_{0,\theta}^2(x,x,l,y)\Big|_{l=x,y=x}
\end{aligned}$$

$$+\frac{1}{2}\sigma(x;\theta)^2\frac{\partial^2}{\partial y^2}g^2_{0,\theta}(x,x,x,y)\Big|_{y=x}+\widetilde{\widetilde{g^2_{0,\theta}}}(x,x,x,x). \tag{10.41}$$

It remains to calculate the derivatives on the right hand side of (10.41). First,

$$\begin{aligned}
&\frac{\partial^2}{\partial h^2}g^2_{0,\theta}(x,h,x,x)\Big|_{h=x}\\
&=2g_{0,\theta}(x,x,x,x)\frac{\partial^2}{\partial h^2}g_{0,\theta}(x,h,x,x)\Big|_{h=x}+2\left(\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\right)^2\\
&=2\left(\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\right)^2.
\end{aligned}\tag{10.42}$$

The last equality in the previous equation (10.42) follows because $g_{0,\theta}(x,x,x,x)=0$, again see (6.54). Analogously, we find

$$\frac{\partial^2}{\partial l^2}g^2_{0,\theta}(x,x,l,x)\Big|_{l=x}=2\left(\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}\right)^2, \tag{10.43}$$

and

$$\frac{\partial^2}{\partial y^2}g^2_{0,\theta}(x,x,x,y)\Big|_{l=y}=2\left(\frac{\partial}{\partial y}g_{0,\theta}(x,x,x,y)\Big|_{y=x}\right)^2. \tag{10.44}$$

For the first cross-term, we have

$$\begin{aligned}
\frac{\partial^2}{\partial h\partial l}g^2_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x} &= 2g_{0,\theta}(x,x,x,x)\frac{\partial^2}{\partial h\partial l}g_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x}\\
&\quad+2\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}\\
&= 2\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}.
\end{aligned}\tag{10.45}$$

And analogously, one obtains

$$\frac{\partial^2}{\partial h\partial y}g^2_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x}=2\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\frac{\partial}{\partial y}g_{0,\theta}(x,x,x,y)\Big|_{y=x}, \tag{10.46}$$

and

$$\frac{\partial^2}{\partial l\partial y}g^2_{0,\theta}(x,h,l,x)\Big|_{h=x,l=x}=2\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}\frac{\partial}{\partial y}g_{0,\theta}(x,x,x,y)\Big|_{y=x}. \tag{10.47}$$

Finally, let us consider the term $\widetilde{\widetilde{g^2_{0,\theta}}}(x,x,x,x)$. By (6.44), we obtain the following equation

$$\begin{aligned}
g^2_{\Delta,\theta}(x,h,l,y)&=\Big(g_{0,\theta}(x,h,l,y)+\sqrt{\Delta}\widetilde{g}_{0,\theta}(x,h,l,y)+\Delta\widetilde{\widetilde{g}}_{0,\theta}(x,h,l,y)+O(\Delta^{3/2})\Big)^2\\
&=g^2_{0,\theta}(x,h,l,y)+2\sqrt{\Delta}g_{0,\theta}(x,h,l,y)\widetilde{g}_{0,\theta}(x,h,l,y)
\end{aligned}$$

$$+\,\Delta\left\{\left(\widetilde{g}_{0,\theta}(x,x,x,x)\right)^2+2g_{0,\theta}(x,x,x,x)\right\}+O(\Delta^{3/2}). \qquad (10.48)$$

By the definitions of the square-root-derivatives in formulae (6.38)-(6.40), the Δ-term in (10.48) corresponds to $\widetilde{\widetilde{g^2}}$, which shows that

$$\widetilde{\widetilde{g^2_{0,\theta}}}(x,x,x,x)=\left(\widetilde{g}_{0,\theta}(x,x,x,x)\right)^2+2g_{0,\theta}(x,x,x,x). \qquad (10.49)$$

By the fact that $g_{0,\theta}(x,x,x,x)=0$ and by the definition of $\mathcal{A}^{(\frac{1}{2})}$, we obtain the final equation

$$\begin{aligned}
&\widetilde{\widetilde{g^2_{0,\theta}}}(x,x,x,x)\\
&=\left(\widetilde{g}_{0,\theta}(x,x,x,x)\right)^2\\
&=\left(-\mathcal{A}^{(\frac{1}{2})}g_{0,\theta}(x,x,x,x)\right)^2\\
&=\left(-\sigma(x;\theta)\sqrt{\frac{2}{\pi}}\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}+\sigma(x;\theta)\sqrt{\frac{2}{\pi}}\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}\right)^2\\
&=\sigma(x;\theta)^2\frac{2}{\pi}\left(\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\right)^2+\sigma(x;\theta)^2\frac{2}{\pi}\left(\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}\right)^2\\
&\quad-2\sigma(x;\theta)^2\frac{2}{\pi}\frac{\partial}{\partial h}g_{0,\theta}(x,h,x,x)\Big|_{h=x}\frac{\partial}{\partial l}g_{0,\theta}(x,x,l,x)\Big|_{l=x}. \qquad (10.50)
\end{aligned}$$

Inserting the above terms for the derivatives of $g^2_{0,\theta}$ and for $\widetilde{\widetilde{g^2_{0,\theta}}}$ into (10.41), we obtain the operator $\mathcal{A}^S$. This completes the proof of the proposition. □

Proof of Proposition 6.3.2.6. For convenience, let us assume that the martingale estimating function depends on h alone. This means, we assume that $g_{\Delta,\theta}$ has the following form

$$g_{\Delta,\theta}(x,h,y)=a(\Delta,x;\theta)\left(\kappa\Big(h-F^H(\Delta,x;\theta)\Big)-\mathbb{E}_{x,\theta}\Big[\kappa\Big(H^Y_\Delta-F^H(\Delta,x;\theta)\Big)\Big]\right). \quad (10.51)$$

In order to satisfy Assumption 6.3.1.3 the function κ must be three times continuously differentiable and κ''' must have polynomial growth near infinity. Condition (6.73) is equivalent to $\kappa'(0)=0$. We expand the expression $\kappa\Big(h-F^H(\Delta,x;\theta)\Big)$ around 0 in order to obtain

$$\begin{aligned}
\kappa\Big(h-F^H(\Delta,x;\theta)\Big)&=\kappa(0)+\frac{1}{2}\kappa''(0)\Big(h-F^H(\Delta,x;\theta)\Big)^2\\
&\quad+\ \frac{1}{6}k'''(\xi)\Big(h-F^H(\Delta,x;\theta)\Big)^3, \qquad (10.52)
\end{aligned}$$

where ξ is between 0 and $h-F^H(\Delta,x;\theta)$. Due to the assumption that κ''' has polynomial growth, we find

$$\mathbb{E}_{x,\theta}\left[\kappa\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)\right] = \kappa(0) + \frac{1}{2}\kappa''(0)\mathbb{E}_{x,\theta}\left[\Big\{H^Y_\Delta - F^H(\Delta,x;\theta)\Big\}^2\right] + O(\Delta^{3/2}). \tag{10.53}$$

By the same argument, an expansion of $g^2_{\Delta,\theta}$ gives the estimate

$$\begin{aligned}&\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}(Y_0,H^Y_\Delta,Y_\Delta)\Big]\\ &= \frac{1}{4}a(\Delta,x;\theta)^2\kappa''(0)\mathbb{E}_{x,\theta}\left[\left\{\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)^2 - \mathbb{E}_{x,\theta}\Big[\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)^2\Big]\right\}^2\right]\\ &\quad + O(\Delta^{5/2}).\end{aligned} \tag{10.54}$$

The statement of Lemma 5.2.1.4 in combination with the scaling property of Brownian motion shows that

$$\begin{aligned}&\mathbb{E}_{x,\theta}\left[\left\{\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)^2 - \mathbb{E}_{x,\theta}\Big[\Big(H_\Delta - F^H(\Delta,x;\theta)\Big)^2\Big]\right\}^2\right]\\ &\qquad = \Delta^2\sigma(x;\theta)^4\mathbb{E}\Big[\Big((H^B_1 - \mathbb{E}H^B_1)^2 - \operatorname{Var}(H^B_1)\Big)^2\Big] + O(\Delta^{5/2}),\end{aligned} \tag{10.55}$$

where, for standard one-dimensional Brownian motion B, H^B_1 denotes the random variable

$$H^B_1 = \sup_{0\le s\le 1} B_s. \tag{10.56}$$

Obviously,

$$\frac{\partial^2}{\partial h^2}g_{0,\theta}(x,h,x)\Big|_{h=x} = a(\Delta,x;\theta)^2\kappa''(0), \tag{10.57}$$

and consequently it remains to state that

$$\begin{aligned}\mathbb{E}\big[((H^B_1 - \mathbb{E}H^B_1)^2 - \operatorname{Var}(H^B_1))^2\big] &= \mathbb{E}\big[(H^B_1 - \mathbb{E}H^B_1)^4\big] - \operatorname{Var}(H^B_1)^2\\ &= 2 - \frac{16}{\pi^2}.\end{aligned} \tag{10.58}$$

Altogether, for the function $g_{\Delta,\theta}$ defined by formula (10.51), we have proved that

$$\begin{aligned}&\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}(Y_0,H^Y_\Delta,Y_\Delta)\Big]\\ &\qquad = \sigma(x;\theta)^4\left(\frac{1}{2} - \frac{4}{\pi^2}\right)\left(\frac{\partial^2}{\partial h^2}g_{0,\theta}(x,h,x)\Big|_{h=x}\right)^2 + O(\Delta^{5/2}).\end{aligned} \tag{10.59}$$

More general functions that depend on both variables h and y, are treated in the same way. One just has to consider all possible partial derivatives of $g_{\Delta,\theta}(x,h,y)$ with respect to h and y separately. The basic ideas behind the proof remain the same as above.

We omit the details here. Instead, we list the remaining relevant moments of (H_1^B, B_1). These moments are easily calculated by means of the joint density of (H_1^B, B_1), which we depicted in (5.49). First, we have

$$\mathbb{E}\big[((B_1 - \mathbb{E}B_1)^2 - \operatorname{Var}(B_1))^2\big] = \mathbb{E}\big[(B_1 - \mathbb{E}B_1)^4\big] - \operatorname{Var}(B_1)^2 = 2, \tag{10.60}$$

and

$$\begin{aligned} &\mathbb{E}\big[((H_1^B - \mathbb{E}H_1^B)(B_1 - \mathbb{E}B_1) - \operatorname{Cov}(H_1^B, B_1))^2\big] \\ &= \mathbb{E}\big[(H_1^B - \mathbb{E}H_1^B)^2(B_1 - \mathbb{E}B_1)^2\big] - \operatorname{Cov}(H_1^B, B_1)^2 = \frac{7}{4} - \frac{10}{3\pi}. \end{aligned} \tag{10.61}$$

Furthermore, the off-diagonal moments are

$$\begin{aligned} &\mathbb{E}\big[((H_1^B - \mathbb{E}H_1^B)^2 - \operatorname{Var}(H_1^B))((H_1^B - \mathbb{E}H_1^B)B_1 - \operatorname{Cov}(H_1^B, B_1))\big] \\ &= \mathbb{E}\big[(H_1^B - \mathbb{E}H_1^B)^3 B_1\big] - \operatorname{Var}(H_1^B)\operatorname{Cov}(H_1^B, B_1) = \frac{7}{4} - \frac{4}{\pi}, \end{aligned} \tag{10.62}$$

$$\begin{aligned} &\mathbb{E}\big[((B_1 - \mathbb{E}B_1)^2 - \operatorname{Var}(B_1))((H_1^B - \mathbb{E}H_t^B)B_1 - \operatorname{Cov}(H_1^B, B_1))\big] \\ &= \mathbb{E}\big[(H_1^B - \mathbb{E}H_1^B)B_1^3\big] - \operatorname{Var}(B_1)\operatorname{Cov}(H_1^B, B_1)\big] = 1, \end{aligned} \tag{10.63}$$

and

$$\begin{aligned} &\mathbb{E}\big[((B_1 - \mathbb{E}B_1)^2 - \operatorname{Var}(B_1))((H_1^B - \mathbb{E}H_1^B)^2 - \operatorname{Var}(H_1^B))\big] \\ &= \mathbb{E}\big[(B_1 - \mathbb{E}B_1)^2(H_1^B - \mathbb{E}H_1^B)^2\big] - \operatorname{Var}(B_1)\operatorname{Var}(H_1^B) = 1 - \frac{4}{3\pi}. \end{aligned} \tag{10.64}$$

Altogether, this proves the result. $\square$

10.3 Proofs of Chapter 7

10.3.1 Proofs of Section 7.4.1 and Section 7.4.2

We give the missing proofs of Theorem 7.4.1.2 and Theorem 7.4.2.2.

Proof of Theorem 7.4.1.2. First, we differentiate (7.255) with respect to t. Due to uniform convergence the result is

$$\begin{aligned}\frac{\partial}{\partial t}v(t,x) =&\sum_{i=1}^{\infty}\epsilon^i\frac{\partial}{\partial t}\phi_i(t,x)\\ =&\epsilon\left\{\frac{y-x}{t}\frac{\partial}{\partial x}\phi_1(t,x)-\frac{1}{2}\beta(x)\right\}\\ &+\sum_{i=2}^{\infty}\epsilon^i\left\{\frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{i-1}(t,x)-\frac{1}{2}\beta(x)\phi_{i-1}(t,x)+\frac{y-x}{t}\frac{\partial}{\partial x}\phi_i(t,x)\right\}.\end{aligned} \tag{10.65}$$

We rearrange the terms on the right hand side of the latter equation and obtain

$$\begin{aligned}\frac{\partial}{\partial t}v(t,x) =&\frac{\epsilon}{2}\sum_{i=2}^{\infty}\epsilon^{i-1}\frac{\partial^2}{\partial x^2}\phi_{i-1}(t,x)+\frac{y-x}{t}\sum_{i=1}^{\infty}\epsilon^i\frac{\partial}{\partial x}\phi_i(t,x)\\ &-\epsilon\left\{\frac{1}{2}\beta(x)-\frac{1}{2}\beta(x)\sum_{i=1}^{\infty}\epsilon^i\phi_i(t,x)\right\}\\ =&\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}v(t,x)+\frac{y-x}{t}\frac{\partial}{\partial x}v(t,x)-\frac{\epsilon}{2}\beta(x)v(t,x).\end{aligned} \tag{10.66}$$

This shows that v satisfies the asserted partial differential equation. The initial condition is verified immediately.

Next, we differentiate ϕ_1^ϵ with respect to t. Due to uniform convergence, the result is

$$\begin{aligned}&\frac{\partial}{\partial t}\phi_1^\epsilon(t,x)\\ =&\sum_{i=1}^{\infty}\epsilon^{i-1}\frac{\partial}{\partial t}\phi_i(t,x)\\ =&\left\{\frac{y-x}{t}\frac{\partial}{\partial x}\phi_1(t,x)-\frac{1}{2}\beta(x)\right\}\\ &+\sum_{i=2}^{\infty}\epsilon^{i-1}\left\{\frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{i-1}(t,x)-\frac{1}{2}\beta(x)\phi_{i-1}(t,x)+\frac{y-x}{t}\frac{\partial}{\partial x}\phi_i(t,x)\right\}.\end{aligned} \tag{10.67}$$

We rearrange the terms on the right hand side of the latter equation and obtain

$$\frac{\partial}{\partial t}\phi_1^\epsilon(t,x)$$

$$
\begin{aligned}
=&\frac{\epsilon}{2}\sum_{i=1}^{\infty}\epsilon^{i-1}\frac{\partial^2}{\partial x^2}\phi_{i-1}(t,x)+\frac{2h-x-y}{t}\sum_{i=1}^{\infty}\epsilon^{i-1}\frac{\partial}{\partial x}\phi_i(t,x)\\
&-\frac{1}{2}\beta(x)-\frac{1}{2}\beta(x)\sum_{i=1}^{\infty}\epsilon^{i}\phi_i(t,x)\\
=&\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}\phi_1^\epsilon(t,x)+\frac{y-x}{t}\frac{\partial}{\partial x}\phi_1^\epsilon(t,x)-\frac{\epsilon}{2}\beta(x)\phi_1^\epsilon(t,x)-\frac{1}{2}\beta(x).
\end{aligned}\tag{10.68}
$$

This shows that $\tilde{\phi}_1^\epsilon$ satisfies the asserted differential equation. Again, the initial condition is trivial to verify. Now, let us differentiate $\phi_k^\epsilon,\ k\geq 2$, with respect to t.

$$
\begin{aligned}
\frac{\partial}{\partial t}\phi_k^\epsilon(t,x)=&\sum_{i=k}^{\infty}\epsilon^{i-k}\frac{\partial}{\partial t}\phi_i(t,x)\\
=&\sum_{i=k}^{\infty}\epsilon^{i-k}\left\{\frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{i-1}(t,x)-\frac{1}{2}\beta(x)\phi_{i-1}(t,x)+\frac{y-x}{t}\frac{\partial}{\partial x}\phi_i(t,x)\right\}.
\end{aligned}\tag{10.69}
$$

Hence, we obtain

$$
\begin{aligned}
\frac{\partial}{\partial t}\phi_k^\epsilon(t,x)=&\frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_{k-1}(t,x)+\frac{\epsilon}{2}\sum_{i=k+1}^{\infty}\epsilon^{i-1-k}\frac{\partial^2}{\partial x^2}\phi_{i-1}(t,x)\\
&-\frac{1}{2}\beta(x)\phi_{k-1}(t,x)-\frac{\epsilon}{2}\beta(x)\sum_{i=k+1}^{\infty}\epsilon^{i-1-k}\phi_{i-1}(t,x)\\
&+\frac{y-x}{t}\sum_{i=k}^{\infty}\epsilon^{i-k}\frac{\partial}{\partial x}\phi_i(t,x)\\
=&\frac{1}{2}\frac{\partial^2}{\partial x^2}\Big\{\phi_{k-1}(t,x)+\epsilon\phi_k^\epsilon(t,x)\Big\}-\frac{1}{2}\beta(x)\Big\{\phi_{k-1}(t,x)+\epsilon\phi_k^\epsilon(t,x)\Big\}\\
&+\frac{y-x}{t}\frac{\partial}{\partial x}\phi_k^\epsilon(t,x).
\end{aligned}\tag{10.70}
$$

By definition, we have

$$
\phi_k^\epsilon(t,x)=\frac{\phi_{k-1}^\epsilon(t,x)-\phi_{k-1}(t,x)}{\epsilon}.\tag{10.71}
$$

Therefore, the previous computations show that ϕ_k^ϵ satisfies the asserted differential equation. The initial condition is trivial to check. And finally, (7.258) holds, due to the definition of $\tilde{\phi}_k^\epsilon$. Recall that the series (7.255) was assumed to converge uniformly. □

Proof of Theorem 7.4.2.2. *(i)* First, we differentiate $\tilde{\phi}_1^\epsilon$, defined by (7.276), with respect to t. Due to uniform convergence the result is

$$
\frac{\partial}{\partial t}\tilde{\phi}_1^\epsilon(t,x)=\sum_{i=1}^{\infty}\epsilon^{i-1}\frac{\partial}{\partial t}\tilde{\phi}_i(t,x)
$$

$$
\begin{aligned}
&= \left\{ \frac{2h-x-y}{t} \frac{\partial}{\partial x} \tilde{\phi}_1(t,x) - \frac{1}{2}\beta(x) \right\} \\
&\quad + \sum_{i=2}^{\infty} \epsilon^{i-1} \left\{ \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{i-1}(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{i-1}(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_i(t,x) \right\}. \qquad (10.72)
\end{aligned}
$$

We rearrange the terms on the right hand side of the latter equation and obtain

$$
\begin{aligned}
\frac{\partial}{\partial t}\tilde{\phi}_1^{\epsilon}(t,x) =& \frac{\epsilon}{2}\sum_{i=1}^{\infty}\epsilon^{i-1}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{i-1}(t,x) + \frac{2h-x-y}{t}\sum_{i=1}^{\infty}\epsilon^{i-1}\frac{\partial}{\partial x}\tilde{\phi}_i(t,x) \\
& - \frac{1}{2}\beta(x) - \frac{1}{2}\beta(x)\sum_{i=1}^{\infty}\epsilon^{i}\tilde{\phi}_i(t,x) \\
=& \frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^{\epsilon}(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^{\epsilon}(t,x) - \frac{\epsilon}{2}\beta(x)\tilde{\phi}_1^{\epsilon}(t,x) - \frac{1}{2}\beta(x). \qquad (10.73)
\end{aligned}
$$

This shows that ϕ_1^{ϵ} satisfies (7.166). The initial and boundary conditions are verified immediately. Now, let us differentiate $\tilde{\phi}_k^{\epsilon},\ k \geq 2$ with respect to t.

$$
\begin{aligned}
&\frac{\partial}{\partial t}\tilde{\phi}_k^{\epsilon}(t,x) \\
&= \sum_{i=k}^{\infty}\epsilon^{i-k}\frac{\partial}{\partial t}\tilde{\phi}_i(t,x) \\
&= \sum_{i=k}^{\infty}\epsilon^{i-k}\left\{ \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{i-1}(t,x) - \frac{1}{2}\beta(x)\tilde{\phi}_{i-1}(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_i(t,x) \right\}. \qquad (10.74)
\end{aligned}
$$

From this expression, one directly obtains

$$
\begin{aligned}
\frac{\partial}{\partial t}\tilde{\phi}_k^{\epsilon}(t,x) =& \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{k-1}(t,x) + \frac{\epsilon}{2}\sum_{i=k+1}^{\infty}\epsilon^{i-1-k}\frac{\partial^2}{\partial x^2}\tilde{\phi}_{i-1}(t,x) \\
& - \frac{1}{2}\beta(x)\tilde{\phi}_{k-1}(t,x) - \frac{\epsilon}{2}\beta(x)\sum_{i=k+1}^{\infty}\epsilon^{i-1-k}\tilde{\phi}_{i-1}(t,x) \\
& + \frac{2h-x-y}{t}\sum_{i=k}^{\infty}\epsilon^{i-k}\frac{\partial}{\partial x}\tilde{\phi}_i(t,x) \\
=& \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big\{\tilde{\phi}_{k-1}(t,x) + \epsilon\tilde{\phi}_k^{\epsilon}(t,x)\Big\} - \frac{1}{2}\beta(x)\Big\{\tilde{\phi}_{k-1}(t,x) + \epsilon\tilde{\phi}_k^{\epsilon}(t,x)\Big\} \\
& + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_k^{\epsilon}(t,x). \qquad (10.75)
\end{aligned}
$$

By definition, we have

$$
\tilde{\phi}_k^{\epsilon}(t,x) = \frac{\tilde{\phi}_{k-1}^{\epsilon}(t,x) - \tilde{\phi}_{k-1}(t,x)}{\epsilon}. \qquad (10.76)
$$

Overall, this shows that ϕ_k^ϵ satisfies the differential equation (7.199). Again, the initial and boundary conditions are trivial to check. Finally, (7.278) holds due to the definition of $\tilde{\phi}_k^\epsilon$. Recall that the participating series' were assumed to converge uniformly.

(ii) The equation of interest is

$$v(t,x) - v_h(t,x) = \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right). \tag{10.77}$$

It is equivalent to (7.279). Let us set $V(t,x) = v(t,x) - v_h(t,x)$. We have to show that the following partial differential equation is satisfied

$$\frac{\partial}{\partial t}V(t,x) = \frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}V(t,x) + \frac{y-x}{t}\frac{\partial}{\partial x}V(t,x) - \frac{\epsilon}{2}\beta(x)V(t,x). \tag{10.78}$$

We differentiate V with respect to t and obtain

$$\begin{aligned}&\frac{\partial}{\partial t}V(t,x)\\ =&\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left\{2\frac{(h-x)(h-y)}{\epsilon t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right) + \epsilon\frac{\partial}{\partial t}\tilde{\phi}_1^\epsilon(t,x)\right\}.\end{aligned} \tag{10.79}$$

We already know from the proof of part (i) that ϕ_1^ϵ satisfies the differential equation (7.166). Therefore,

$$\begin{aligned}&\frac{\partial}{\partial t}V(t,x)\\ =&\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)2\frac{(h-x)(h-y)}{\epsilon t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\\ &+\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\epsilon\left\{\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x) + \frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\right.\\ &\left.\qquad -\frac{\epsilon}{2}\beta(x)\tilde{\phi}_1^\epsilon(t,x) - \frac{1}{2}\beta(x)\right\}.\end{aligned} \tag{10.80}$$

Furthermore, by differentiating V once and twice with respect to x, one obtains the two equations

$$\begin{aligned}&\frac{\partial}{\partial x}V(t,x)\\ =&\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left\{2\frac{h-y}{\epsilon t}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right) + \epsilon\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\right\}\end{aligned} \tag{10.81}$$

and

$$\frac{\partial^2}{\partial x^2}V(t,x)$$

$$
\begin{aligned}
&= \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\left\{4\frac{(h-y)^2}{\epsilon^2 t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)+4\frac{h-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\right\}\\
&\quad + \exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\epsilon\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x). \qquad (10.82)
\end{aligned}
$$

We solve the latter two equations for $\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)$ and $\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x)$ in order to obtain

$$
\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x) = \frac{1}{\epsilon}\left\{\exp\left(2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{\partial}{\partial x}V(t,x) - 2\frac{h-y}{\epsilon t}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\right\} \qquad (10.83)
$$

and

$$
\begin{aligned}
&\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x)\\
&=\frac{1}{\epsilon}\left\{\exp\left(2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{\partial^2}{\partial x^2}V(t,x) - 4\frac{(h-y)^2}{\epsilon^2 t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\right.\\
&\quad\left. - 4\frac{h-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\right\}\\
&=\frac{1}{\epsilon}\left\{\exp\left(2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{\partial^2}{\partial x^2}V(t,x) - 4\frac{(h-y)^2}{\epsilon^2 t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\right.\\
&\quad\left. - 4\frac{h-y}{\epsilon t}\left[\exp\left(2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{\partial}{\partial x}V(t,x) - 2\frac{h-y}{\epsilon t}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\right]\right\}.
\end{aligned}
\qquad (10.84)
$$

Furthermore,

$$
\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right) = \exp\left(2\frac{(h-x)(h-y)}{\epsilon t}\right)V(t,x). \qquad (10.85)
$$

Inserting (10.85) into (10.80), we obtain

$$
\begin{aligned}
&\frac{\partial}{\partial t}V(t,x)\\
&=\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)2\frac{(h-x)(h-y)}{\epsilon t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\\
&\quad+\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\epsilon\left\{\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}\tilde{\phi}_1^\epsilon(t,x)+\frac{2h-x-y}{t}\frac{\partial}{\partial x}\tilde{\phi}_1^\epsilon(t,x)\right\}\\
&\quad-\frac{\epsilon}{2}\beta(x)V(t,x) \qquad (10.86)
\end{aligned}
$$

Next, we plug (10.83) and (10.84) into the previous equation, which yields

$$
\begin{aligned}
\frac{\partial}{\partial t}V(t,x) =& 2\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{(h-x)(h-y)}{\epsilon t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\\
&+\frac{2h-x-y}{t}\frac{\partial}{\partial x}V(t,x)
\end{aligned}
$$

$$
\begin{aligned}
&-2\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{2h-x-y}{t}\frac{h-y}{\epsilon t}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\\
&+\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}V(t,x)\\
&-2\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{(h-y)^2}{\epsilon t^2}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\\
&-\frac{2(h-y)}{t}\frac{\partial}{\partial x}V(t,x)\\
&+4\exp\left(-2\frac{(h-x)(h-y)}{\epsilon t}\right)\frac{(h-y)^2}{\epsilon t}\left(1+\epsilon\tilde{\phi}_1^\epsilon(t,x)\right)\\
&-\frac{\epsilon}{2}\beta(x)V(t,x)\\
=&\frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}V(t,x)+\frac{2h-x-y}{t}\frac{\partial}{\partial x}V(t,x)-\frac{2(h-y)}{t}\frac{\partial}{\partial x}V(t,x)\\
&-\frac{\epsilon}{2}\beta(x)V(t,x).
\end{aligned}
\tag{10.87}
$$

Since $(2h-x-y)-2(h-y)=y-x$, this gives the asserted differential equation. The boundary condition trivially holds. □

10.3.2 Proofs of Section 7.4.3

Proof of the convergence criterion - case without boundary conditions

In the sequel, we will prove the following proposition.

Proposition 10.3.2.1. *Let ϕ_1 be the function given in* (7.246) *and let the functions ϕ_k be defined by the recursion* (7.247), (7.249). *If $\beta(x)=x^2$ or $-x^2$, then the functions ϕ_k satisfy for all $t\geq 0$ and for all $x,y\in\mathbb{R}$ the estimate*

$$
\left|\phi_k(t,x,y)\right|\leq 2^k t^k\exp\left(|x|^2\vee|y|^2\right),\quad k\in\mathbb{N}. \tag{10.88}
$$

Before being able to prove this result, we have to state some auxiliary lemmata.

Lemma 10.3.2.2. *Let $k,m\in\mathbb{N}$ and $t>0$. For $x,y\in\mathbb{R}$, the integral*

$$
\int_0^t\left(x+\frac{s}{t}(y-x)\right)^m(t-s)^k ds \tag{10.89}
$$

is a polynomial of the following form

$$
\frac{t^{k+1}}{k+1}\sum_{i=0}^m c_i x^i y^{m-i}. \tag{10.90}
$$

Moreover, the coefficients c_i are all non-negative and satisfy $\sum_{i=0}^m c_i=1$.

Proof. First, let us take into account that

$$\left(x + \frac{s}{t}(y-x)\right)^m = \left(\left(1 - \frac{s}{t}\right)x + \frac{s}{t}y\right)^m. \tag{10.91}$$

For $0 \le i \le m$, the integrals

$$\int_0^t (t-s)^k \left(1 - \frac{s}{t}\right)^i \left(\frac{s}{t}\right)^{m-i} ds \tag{10.92}$$

are obviously positive and it is not difficult to verify that (10.92) equals

$$t^{k+1}\frac{(k+i)!(m-i)!}{(m+k+1)!}. \tag{10.93}$$

Hence, the integral (10.89) has the representation (10.90) with non-negative coefficients c_i. In order to prove the remaining property set $x = y$. In this case, the integral (10.89) becomes

$$x^m \int_0^t (t-s)^k ds = \frac{1}{k+1} t^{k+1} x^m. \tag{10.94}$$

Therefore $\sum_{i=0}^m c_i = 1$, and the assertion follows. □

Lemma 10.3.2.3. *Let ϕ_1 be the function given in* (7.246) *and let the functions ϕ_k be defined by the recursion* (7.247), (7.249). *If $\beta(x) = x^2$ or $-x^2$, then ϕ_k consists of 2^k (or less) polynomials, each of which has the form*

$$t^k \sum_{j=0}^m b_j x^j y^{m-j}. \tag{10.95}$$

Here, $m \le 2 \cdot k$ and either $b_j \ge 0,\ \forall j = 0, \dots, m$ or $b_j \le 0,\ \forall j = 0, \dots, m$. Moreover, the coefficients satisfy

$$\left|\sum_{j=0}^m b_j\right| = \sum_{j=0}^m |b_j| \le \frac{1}{m!!}. \tag{10.96}$$

For $m \in \mathbb{N}$, the expression $m!!$ denotes the double factorial of m. This means $m!! = m(m-2) \cdot \ldots \cdot 4 \cdot 2$ if m is even, and $m!! = m(m-2) \cdot \ldots \cdot 3 \cdot 1$ if m is odd.

Proof. First note that, for $\beta(x) = x^2$,

$$\phi_1(t,x,y) = \frac{1}{2}\int_0^t \beta\left(x + \frac{u}{t}(y-x)\right) ds = \frac{1}{6} t\left(x^2 + yx + y^2\right), \tag{10.97}$$

and, for $\beta(x) = -x^2$,

$$\phi_1(t,x,y) = \frac{1}{2}\int_0^t \beta\left(x + \frac{u}{t}(y-x)\right) ds = -\frac{1}{6} t\left(x^2 + yx + y^2\right). \tag{10.98}$$

In the sequel we will only consider the case $\beta(x) = x^2$. The proof for the case $\beta(x) = -x^2$

works in exactly the same way. Now, assume that, for $k > 1$, the function ϕ_k is given by

$$\phi_k(t, x, y) = \sum_{n=0}^{2^k} \vartheta_n(t, x, y) \tag{10.99}$$

where each ϑ_n is a polynomial of the form (10.95). We apply the differential operator $\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}\beta(\cdot)$ to the polynomial (10.95) in order to obtain

$$\frac{1}{2}t^k \sum_{j=2}^{m} b_j j(j-1)x^{j-2}y^{m-j} + \frac{1}{2}t^k \sum_{j=0}^{m} b_j x^{j+2}y^{m-j}. \tag{10.100}$$

Recall the definition of the recursion (7.247), (7.249). We replace t by $t-s$ and x by $\left(x + \frac{s}{t}(y-x)\right)$ in the latter expression and integrate from 0 to t with respect to s. The resulting expression is

$$\begin{aligned} &\frac{1}{2}\sum_{j=2}^{m} b_j j(j-1)\left\{\int_0^t (t-s)^k \left(x + \frac{s}{t}(y-x)\right)^{j-2} ds\right\} y^{m-j} \\ &\qquad + \frac{1}{2}\sum_{j=0}^{m} b_j \left\{\int_0^t (t-s)^k \left(x + \frac{s}{t}(y-x)\right)^{j+2} ds\right\} y^{m-j}. \end{aligned} \tag{10.101}$$

We consider both terms individually. By the result of Lemma 10.3.2.2, the first term in (10.101) becomes

$$\begin{aligned} &\frac{1}{2}\sum_{j=2}^{m} b_j j(j-1)\left\{\int_0^t (t-s)^k \left(x + \frac{s}{t}(y-x)\right)^{j-2} ds\right\} y^{m-j} \\ &\qquad = \frac{t^{k+1}}{2(k+1)} \sum_{j=2}^{m} b_j j(j-1)\left\{\sum_{n=0}^{j-2} c_{j,n} x^n y^{j-2-n}\right\} y^{m-j}, \end{aligned} \tag{10.102}$$

where the $c_{j,n}$ are all positive and satisfy $\sum_{n=0}^{j-2} c_{j,n} = 1$ for all $j = 2, \ldots, m$. Due to our assumptions, the b_j are either all positive or all negative. Since $m \leq 2k$, we obtain

$$\begin{aligned} &\frac{1}{2}\sum_{j=2}^{m} b_j j(j-1)\left\{\int_0^t (t-s)^k \left(x + \frac{s}{t}(y-x)\right)^{j-2} ds\right\} y^{m-j} \\ &= t^{k+1}\sum_{j=2}^{m} b_j \frac{j(j-1)}{2(k+1)}\left\{\sum_{n=0}^{j-2} c_{j,n} x^n y^{j-2-n}\right\} y^{m-j} \\ &= t^{k+1}\sum_{j=0}^{m-2} d_j x^j y^{m-2-j}, \end{aligned} \tag{10.103}$$

where the new coefficients d_j are either all positive or all negative and by setting $x = y$

we see that they must satisfy

$$\left|\sum_{j=0}^{m-2} d_j\right| \le \frac{m(m-1)}{2\cdot(k+1)} \sum_{j=2}^{m} |b_j| \le \frac{m}{m!!} = \frac{1}{(m-2)!!}. \tag{10.104}$$

We have to consider the second term in (10.101). By Lemma 10.3.2.2 we obtain

$$\begin{aligned}\frac{1}{2(k+1)} \sum_{j=0}^{m} b_j \left\{ \int_0^t (t-s)^k \left(x + \frac{s}{t}(y-x)\right)^{j+2} ds \right\} y^{m-j} \\ = \frac{1}{2(k+1)} \sum_{j=0}^{m} b_j \left\{ \sum_{n=0}^{j+2} \tilde{c}_{j,n} x^n y^{j+2-n} \right\} y^{m-j},\end{aligned} \tag{10.105}$$

where the $\tilde{c}_{j,n}$ are all positive and satisfy $\sum_{n=0}^{j+2} \tilde{c}_{j,n} = 1$ for all $j = 0, \dots, m$. Recall that, due to the assumption, the b_j all have the same sign. Since $2(k+1) \ge m+2$, we obtain

$$\begin{aligned}&\frac{1}{2(k+1)} \sum_{j=0}^{m} b_j \left\{ \int_0^t (t-s)^k \left(x + \frac{s}{t}(y-x)\right)^{j+2} ds \right\} y^{m-j} \\ &= t^{k+1} \sum_{j=0}^{m+2} \tilde{d}_j x^j y^{m+2-j},\end{aligned} \tag{10.106}$$

where the new coefficients $\tilde{d}_j$ all have the same sign and, by setting $x = y$, we see that they must satisfy

$$\left|\sum_{j=0}^{m+2} \tilde{d}_j\right| \le \frac{1}{2(k+1)} \sum_{j=0}^{m} |b_j| \le \frac{1}{(m+2)\cdot m!!} = \frac{1}{(m+2)!!}. \tag{10.107}$$

Thus, we have shown that, if we apply the recursion formula (7.247), (7.249) to one of the functions ϑ_n in (10.99), then the results are two polynomials of the following form

$$t^k \sum_{j=0}^{m} d_j x^j y^{m-j}, \tag{10.108}$$

where $m \le 2(k+1)$ and the coefficients d_j satisfy

$$\left|\sum_{j=0}^{m} d_j\right| = \sum_{j=0}^{m} |d_j| \le \frac{1}{m!!}. \tag{10.109}$$

Therefore, ϕ_{k+1} satisfies the asserted property and we are able to conclude by induction. □

We are now able to proof the above proposition.

Proof (of Proposition 10.3.2.1). According to Lemma 10.3.2.3, the function ϕ_k is given by

$$\phi_k(t,x,y) = \sum_{n=0}^{2^k} \vartheta_n(t,x,y), \tag{10.110}$$

where each function ϑ_n is a polynomial of the form

$$\vartheta_n(t,x,y) = t^k \sum_{i=0}^{\alpha_n} b_{n,i} x^i y^{\alpha_n - 1}. \tag{10.111}$$

Here, $\alpha_n \in \mathbb{N}$, $\alpha_n \leq 2k$ and the coefficients $b_{n,i}$ satisfy

$$\left| \sum_{i=0}^{\alpha_n} b_{n,i} \right| = \sum_{i=0}^{\alpha_n} |b_{n,i}| \leq \frac{1}{\alpha_n!!}, \tag{10.112}$$

for all $n = 0, \ldots, 2^k$. Particularly, for $x = y$, we find the estimate

$$|\vartheta_n(t,x,x)| = |x|^{\alpha_n} t^k \sum_{i=0}^{\alpha_n} b_{n,i} \leq 2^k |x|^{\alpha_n} t^k \frac{1}{\alpha_n!!} \tag{10.113}$$

And therefore, we conclude that

$$\left|\phi_k(t,x,y)\right| = \sum_{n=0}^{2^k} \left|\vartheta_n(t,x,y)\right| \leq t^k \sum_{n=0}^{2^k} \Big(|x| \vee |y|\Big)^{\alpha_n} \frac{1}{\alpha_n!!} \tag{10.114}$$

Note that we did not determine the exact values of the α_n. But, since $\alpha_n \leq 2k$, we have the estimate

$$\sum_{n=0}^{2^k} \Big(|x| \vee |y|\Big)^{\alpha_n} \frac{1}{\alpha_n!!} \leq 2^k \sum_{j=0}^{2^k} \Big(|x|^2 \vee |y|^2\Big)^j \frac{1}{j!!}. \tag{10.115}$$

Consequently, we are able to choose an exponential series as an upper bound in order to obtain

$$\left|\phi_k(t,x,y)\right| \leq 2^k t^k \exp\Big(|x|^2 \vee |y|^2\Big). \tag{10.116}$$

This shows the result. □

Proof of the convergence criterion - case with boundary conditions

We will prove the following proposition.

Proposition 10.3.2.4. *For $k \in \mathbb{N}$, let $\tilde{\phi}_k$ be the functions defined in Definition 7.4.0.12 and let $\beta(x) = x^2$ or $-x^2$. Then, for all $x, h, y \in \mathbb{R}$ with $x, y \leq h$,*

$$\left|\tilde{\phi}_k(t,x,h,y)\right| \leq 10^k t^k \exp\left(\Big\{|2h - x|^2 \vee |y|^2\Big\} + \Big\{|2h - y|^2 \vee |x|^2\Big\}\right). \tag{10.117}$$

Before we are able to prove this proposition, some auxiliary results are necessary. The ideas behind the proofs are mainly the same as for the case without boundary conditions. Therefore, we will not carry everything through in detail.

Lemma 10.3.2.5. *Let $k, l, m \in \mathbb{N}$. Then, for $x, y \in \mathbb{R}$,*

$$\begin{aligned}
&\int_0^t (t-s)^k \left(2h - x - \frac{s}{t}(2h-x-y)\right)^l \left(x + \frac{s}{t}(2h-x-y)\right)^m \\
&= \frac{t^{k+1}}{k+1} \sum_{i=0}^{l} \sum_{j=0}^{m} c_{i,j} (2h-x)^i y^{l-i} (2h-y)^j x^{m-j} \qquad (10.118)
\end{aligned}$$

where the coefficients satisfy $c_{i,j} \geq 0$ for all $i = 0, \ldots, l$, and for all $j = 0, \ldots, m$. Moreover, $\sum_{i=0}^{l} \sum_{j=0}^{m} c_{i,j} = 1$.

Proof. First let us note that

$$\begin{aligned}
&\left(2h - x - \frac{s}{t}(2h-x-y)\right)^l \left(x + \frac{s}{t}(2h-x-y)\right)^m \\
&= \left(\left(1 - \frac{s}{t}\right)(2h-x) + \frac{s}{t} y\right)^l \left(\left(1 - \frac{s}{t}\right) x + \frac{s}{t}(2h-y)\right)^m . \qquad (10.119)
\end{aligned}$$

For $0 \leq i \leq l$ and $0 \leq j \leq m$, the integrals

$$\int_0^t (t-s)^k \left(1 - \frac{s}{t}\right)^i \left(\frac{s}{t}\right)^{l-i} \left(1 - \frac{s}{t}\right)^j \left(\frac{s}{t}\right)^{m-j} ds \qquad (10.120)$$

are clearly positive. Indeed, it can be shown that (10.120) equals

$$t^{k+1} \frac{(k+i+j)!(l+m-i-j)!}{(k+l+m+2)!}. \qquad (10.121)$$

Therefore, we proved the representation (10.118). And clearly, all the coefficients $c_{i,j}$ are non-negative. Finally, the relation $\sum_{i=0}^{l} \sum_{j=0}^{m} c_{i,j} = 1$ follows by setting $h = x = y$. In this case the formula (10.118) shrinks to

$$x^{m+l} \int_0^t (t-s)^k ds = x^{m+l} \frac{t^{k+1}}{k+1}. \qquad (10.122)$$

Altogether, this shows the result. □

We need to introduce an auxiliary sequence of functions. The functions will be described in the following definition.

Definition 10.3.2.6. Let $\beta \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ and, for $x, h, y \in \mathbb{R}$ with $x, y < h$, set

$$\tilde{\tilde{\phi}}_1(t, x, h, y) = \int_0^t \beta(\bar{\rho}_s) ds + \int_0^t \beta(\bar{\bar{\rho}}_s) ds, \qquad (10.123)$$

where $\bar{\rho}_s = \bar{\rho}_s^{(t,x,h,y)} = x + \frac{s}{t}(2h - x - y)$ and $\bar{\bar{\rho}}_s = \bar{\bar{\rho}}_s^{(t,x,h,y)} = 2h - x - \frac{s}{t}(2h - x - y)$ for $s \in [0, t]$. For $k \geq 2$, recursively define

$$\tilde{\tilde{g}}_{k-1}(t, x, h, y) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\tilde{\tilde{\phi}}_{k-1}(t, x, h, y) - \frac{1}{2}\beta(x)\tilde{\tilde{\phi}}_{k-1}(t, x, h, y) \tag{10.124}$$

and

$$\tilde{\tilde{\phi}}_k(t, x, h, y) = \int_0^t \tilde{\tilde{g}}_{k-1}(t - s, \bar{\rho}_s, h, y)ds + \int_0^t \tilde{\tilde{g}}_{k-1}(t - s, \bar{\bar{\rho}}_s, h, y)ds. \tag{10.125}$$

Remark 10.3.2.7. Note, that the role of the paths $\bar{\rho}.$ and $\bar{\bar{\rho}}.$ was already discussed in Remark 7.3.2.11.

Lemma 10.3.2.8. *For $k \in \mathbb{N}$, the function $\tilde{\tilde{\phi}}_k(t, x, h, y)$ consists of 10^k (or less) polynomials of the form*

$$t^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}(2h - x)^i y^{l-i}(2h - y)^{m-j}x^j \tag{10.126}$$

where $m, l \leq 2k$ and where the coefficients $b_{i,j}$ are either all positive or all negative and satisfy $\sum_{i=0}^{l}\sum_{j=0}^{m}|b_{i,j}| \leq \frac{1}{l!!\cdot m!!}$.

In order to prove Lemma 10.3.2.8, we need another auxiliary result.

Lemma 10.3.2.9. *Let $m \in \mathbb{N}$. Then*

$$\frac{(2m)!!}{(2m-1)!!} \sim \sqrt{\pi m} \tag{10.127}$$

and

$$\frac{(2m+1)!!}{(2m)!!} \sim \frac{2}{\sqrt{\pi}}\sqrt{m}. \tag{10.128}$$

Proof. It follows by direct calculations that

$$(2m)!! = 2^m m! \quad \text{and} \quad (2m-1)!! = \frac{(2m)!}{2^m m!}. \tag{10.129}$$

Particularly, $(2m+1)!! = \frac{(2m+2)!}{2^{m+1}(m+1)!}$. Stirling's formula tells us that

$$m! \sim \sqrt{2\pi m}\left(\frac{m}{e}\right)^m. \tag{10.130}$$

Formula (10.127) follows directly. Straightforward calculations show that

$$\frac{(2m+1)!!}{(2m)!!} \sim 2\sqrt{\frac{m}{\pi}}\frac{1}{e}\left(1 + \frac{1}{m}\right)^{m+1}. \tag{10.131}$$

Clearly, $(1 + 1/m)^m \to e$, as $m \to \infty$, and (10.128) follows. $\square$

Proof (of Lemma 10.3.2.8). It is trivial to see that the assertion holds for $\tilde{\tilde{\phi}}_1$ for both the case $\beta(x) = x^2$ and the case $\beta(x) = -x^2$. In the sequel we will only consider $\beta(x) = x^2$. The case $\beta(x) = -x^2$ follows in an analogous way. Let us note that for a polynomial $\mathbb{p}(x) = (2h-x)^i x^j$ the following two relations are satisfied

$$\mathbb{p}(\bar{\rho}_s) = \left(2h - x - \frac{s}{t}(2h-x-y)\right)^i \left(x + \frac{s}{t}(2h-x-y)\right)^j = \bar{\bar{\rho}}_s^i \cdot \bar{\rho}_s^j \tag{10.132}$$

and

$$\mathbb{p}(\bar{\bar{\rho}}_s) = \left(x + \frac{s}{t}(2h-x-y)\right)^i \left(2h - x - \frac{s}{t}(2h-x-y)\right)^j = \bar{\rho}_s^i \cdot \bar{\bar{\rho}}_s^j. \tag{10.133}$$

Moreover, it is easy to show that

$$\frac{1}{2}\frac{\partial^2}{\partial x^2}\mathbb{p}(x) = \frac{1}{2}\Big\{j(j-1)(2h-x)^i x^{j-2} - 2ij(2h-x)^{i-1}x^{j-1} + i(i-1)(2h-x)^{i-2}x^j\Big\}. \tag{10.134}$$

and

$$\frac{1}{2}x^2\mathbb{p}(x)^2 = (2h-x)^i x^{j+2}. \tag{10.135}$$

We see that an application of the operator $\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}\beta(\cdot)$ to the function

$$\vartheta(t,x) = t^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}(2h-x)^i y^{l-i}(2h-y)^{m-j}x^j, \tag{10.136}$$

which corresponds to a function of the type (10.126), results in 5 polynomials. Each of these polynomials has then to be evaluated once at $(t-s, \bar{\rho}_s)$ and once at $(t-s, \bar{\bar{\rho}}_s)$. Concretely, the function

$$\begin{aligned} g_{aux}(s,x) = &\left\{\frac{1}{2}\frac{\partial^2}{\partial y^2}\vartheta(u,y) - \frac{1}{2}\beta(y)\vartheta(u,y)\right\}\Big|_{(u,y)=\left(t-s,\bar{\rho}_s\right)} \\ &+ \left\{\frac{1}{2}\frac{\partial^2}{\partial y^2}\vartheta(u,y) - \frac{1}{2}\beta(y)\vartheta(u,y)\right\}\Big|_{(u,y)=\left(t-s,\bar{\bar{\rho}}_s\right)} \end{aligned} \tag{10.137}$$

consists of the following 8 polynomials

$$\frac{1}{2}(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}i(i-1)\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{i-2} y^{l-i}(2h-y)^{m-j}\left(x+\frac{s}{t}(2h-x-y)\right)^j$$

$$-(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}ij\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{i-1} y^{l-i}(2h-y)^{m-j}\left(x+\frac{s}{t}(2h-x-y)\right)^{j-1}$$

$$\frac{1}{2}(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}j(j-1)\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{i} y^{l-i}(2h-y)^{m-j}\left(x+\frac{s}{t}(2h-x-y)\right)^{j-2}$$

$$
\begin{aligned}
&-\frac{1}{2}(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}\left(2h-x-\frac{s}{t}(2h-x-y)\right)^i y^{l-i}(2h-y)^{m-j}\left(x+\frac{s}{t}(2h-x-y)\right)^{j+2}\\
&\frac{1}{2}(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}i(i-1)\left(x+\frac{s}{t}(2h-x-y)\right)^{i-2} y^{l-i}(2h-y)^{m-j}\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{j}\\
&-(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}ij\left(x+\frac{s}{t}(2h-x-y)\right)^{i-1} y^{l-i}(2h-y)^{m-j}\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{j-1}\\
&\frac{1}{2}(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}j(j-1)\left(x+\frac{s}{t}(2h-x-y)\right)^{i} y^{l-i}(2h-y)^{m-j}\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{j-2}\\
&-\frac{1}{2}(t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}\left(x+\frac{s}{t}(2h-x-y)\right)^{i} y^{l-i}(2h-y)^{m-j}\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{j+2}.
\end{aligned}
\tag{10.138}
$$

We have to be careful, since two of the above polynomials have the additional factor 2. Thus, effectively we have 10 polynomials. The definition of $\tilde{\tilde{\phi}}_{k+1}$ requires to calculate

$$\tilde{\tilde{\phi}}_{k+1}(t,x) = \int_0^t g_{aux}(s,x)ds. \tag{10.139}$$

Hence, we have to integrate each of the polynomials in (10.138) from 0 to t with respect to s. The results are 10 polynomials of the form (10.126) with k replaced by $k+1$. This can be shown by means of Lemma 10.3.2.5. We will prove this fact only for the second polynomial in (10.138). By Lemma 10.3.2.5, we have

$$
\begin{aligned}
&\frac{1}{2}\int_0^t (t-s)^k \sum_{i=0}^{l}\sum_{j=0}^{m} b_{i,j}ij\left(2h-x-\frac{s}{t}(2h-x-y)\right)^{i-1} y^{l-i}(2h-y)^{m-j}\left(x+\frac{s}{t}(2h-x-y)\right)^{j-1} ds\\
&=\frac{1}{2}\frac{t^{k+1}}{k+1}\sum_{i=1}^{l}\sum_{j=1}^{m} b_{i,j}ij\sum_{\alpha_i=0}^{i}\sum_{\alpha_j=0}^{j} c^{(i,j)}_{\alpha_i,\alpha_j}(2h-x)^{\alpha_i}y^{i-\alpha_i}(2h-y)^{\alpha_j}x^{j-\alpha_j},
\end{aligned}
\tag{10.140}
$$

where the $c^{(i,j)}_{\alpha_i,\alpha_j}$ are all positive and satisfy $\sum_{\alpha_i=0}^{i}\sum_{\alpha_j=0}^{j} c^{(i,j)}_{\alpha_i,\alpha_j} = 1$, for all $i = 1, \ldots, l$, $j = 1, \ldots, m$. By assumption, all the coefficients $b_{i,j}$ have the same sign, and thus, by rearranging the terms, it follows that the right hand side of (10.140) equals

$$t^k \sum_{i=0}^{l-1}\sum_{j=0}^{m-1} d_{i,j}(2h-x)^i y^{l-1-i}(2h-y)^j x^{m-1-j}, \tag{10.141}$$

where all the new coefficients $d_{i,j}$ have the same sign. Moreover, setting $h = x = y$, we see that the $d_{i,j}$ satisfy

$$\sum_{i=0}^{l-1}\sum_{j=0}^{m-1} |d_{i,j}| = \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{m}\frac{ij}{k+1}|b_{i,j}| \le \frac{1}{2}\frac{ml}{k+1}\sum_{j=1}^{m}|b_{i,j}| \le \frac{1}{2}\frac{ml}{k+1}\frac{1}{m!!\cdot l!!}. \tag{10.142}$$

By Lemma 10.3.2.9 the right hand side of the previous equation can be bounded in the following way

$$\begin{aligned}\frac{ml}{2(k+1)}\frac{1}{m!!\cdot l!!} &= \frac{1}{2(k+1)}\frac{1}{(m-2)!!\cdot(l-2)!!}\\ &\leq \frac{\sqrt{m}\sqrt{l}}{2(k+1)}\frac{1}{(m-1)!!\cdot(l-1)!!}\\ &\leq \frac{1}{(m-1)!!\cdot(l-1)!!},\end{aligned} \tag{10.143}$$

The last inequality follows because $m, l \leq 2k$. And thus, we have the overall estimate

$$\sum_{i=0}^{l-1}\sum_{j=0}^{m-1}|d_{i,j}| \leq \frac{1}{(m-1)!!\cdot(l-1)!!}. \tag{10.144}$$

This shows that, integrating the second polynomial in (10.138), we obtain a term that has the asserted property for $k+1$. The other terms in (10.138) are even simpler to estimate. The proof works exactly along the same lines as the proof of Lemma 10.3.2.3, where we considered the case with no boundary condition. Note again, that the factor $1/2$ is consumed during the integration. A concise outline of these results involves a lot of tedious notations. For convenience we omit further details here. All in all, the proof follows by induction, since an application of the recursion formula given in Definition 10.3.2.6 to a function of the type (10.126) yields a function of the same type with k replaced by $k+1$. □

Proof (of Proposition 10.3.2.4). For $\beta : \mathbb{R} \to \mathbb{R}$, for $t > 0$ and $x, h, y \in \mathbb{R}$, $x, y < h$, the function $\tilde{\phi}_1(t,x,h,y)$ is given by

$$\begin{aligned}&\tilde{\phi}_1(t,x,h,y)\\ &= \int_0^t \beta(\rho_s)ds\\ &= \int_0^{t\frac{h-x}{2h-x-y}} \beta\left(x+\frac{s}{t}(2h-x-y)\right)ds + \int_{t\frac{h-x}{2h-x-y}}^t \beta\left(x+\frac{t-s}{t}(2h-x-y)\right)ds.\end{aligned} \tag{10.145}$$

This expression is contained in the expression $\tilde{\tilde{\phi}}_1$, which is a sum of two integrals along linear paths. Concretely,

$$\begin{aligned}&\tilde{\tilde{\phi}}_1(t,x,h,y)\\ &= \int_0^t g\left(x+\frac{s}{t}(2h-x-y)\right)ds + \int_0^t g\left(2h-x+\frac{s}{t}(2h-x-y)\right)ds.\end{aligned} \tag{10.146}$$

Note that the path ρ_s coincides with the path $\bar{\rho}_s$ if $s \in [0, t(h-x)/(2h-x-y)]$ and with the path $\bar{\bar{\rho}}_s$ on the remaining time interval, i.e. if $s \in [t(h-x)/(2h-x-y), t]$. We can sum up our deliberations in the following way:

$$\tilde{\tilde{\phi}}_1(t,x,h,y) = \tilde{\phi}_1(t,x,h,y) + \text{"two integral terms"}. \tag{10.147}$$

By the definition of the recursion that defines the functions $\tilde{\phi}_k$, it is easy to see that also

$$\tilde{\tilde{\phi}}_k(t,x,h,y) = \tilde{\phi}_k(t,x,h,y) + \text{"some integral terms"}, \tag{10.148}$$

for $k \geq 1$. We already mentioned this fact in Remark 7.3.2.11. The "formula" (10.148) holds, since in the recursion step, the integrals

$$\tilde{\phi}_k(t,x) = \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{k-1}(t-s,\bar{\rho}_s)ds + \int_{t\frac{h-x}{2h-x-y}}^t g_{k-1}(t-s,\bar{\bar{\rho}}_s)ds, \tag{10.149}$$

that define the $\tilde{\phi}_k$, are augmented in the following way

$$\begin{aligned}
&\int_0^t \tilde{g}_{k-1}(t-s,\bar{\rho}_s)ds + \int_0^t g_{k-1}(t-s,\bar{\bar{\rho}}_s)ds \\
&= \int_0^{t\frac{h-x}{2h-x-y}} \tilde{g}_{k-1}(t-s,\bar{\rho}_s)ds + \int_{t\frac{h-x}{2h-x-y}}^t g_{k-1}(t-s,\bar{\bar{\rho}}_s)ds + \text{"two integral terms"} \\
&= \tilde{\phi}_k(t,s) + \text{"two integral terms"}.
\end{aligned} \tag{10.150}$$

Note again, that ρ_s coincides with $\bar{\rho}_s$ on $[0, t(h-x)/(2h-x-y)]$ and with $\bar{\bar{\rho}}_s$ on $[t(h-x)/(2h-x-y), t]$. This explains (10.148). We see that a bound for $\tilde{\tilde{\phi}}_k$ also bounds $\tilde{\phi}_k$. For $\beta(x) = x^2$ or $-x^2$ the specific structure of the functions $\tilde{\tilde{\phi}}_k$ was determined in Lemma 10.3.2.8. We find the upper bound

$$\left|\tilde{\tilde{\phi}}_k(t,x,h,y)\right| \leq t^k \sum_{i=0}^{10^k} \frac{\left(|2h-x|^{a_i^{(1)}} \vee |y|^{a_i^{(1)}}\right) \cdot \left(|2h-y|^{a_i^{(1)}} \vee |x|^{a_i^{(1)}}\right)}{a_i^{(1)}!! \cdot a_i^{(2)}!!}, \tag{10.151}$$

where the $a_i^{(1)}, a_i^{(2)} \leq 2k$. We did not calculate the integers $a_i^{(1)}, a_i^{(2)}$, but we can simply bound each of the 10^k addends in (10.151) by a product of two exponential series, which results in the estimate

$$\left|\tilde{\tilde{\phi}}_k(t,x,h,y)\right| \leq 10^k t^k \exp\left(\left\{|2h-x|^2 \vee |y|^2\right\} + \left\{|2h-y|^2 \vee |x|^2\right\}\right). \tag{10.152}$$

Recall, that an upper bound for $\tilde{\tilde{\phi}}_k$ is also an upper bound for $\tilde{\phi}_k$. Altogether, this shows the result. □

10.4 Proofs of Chapter 8

10.4.1 Table of coefficients belonging to the Hermite polynomials

Recall the definition of the function Ξ in (8.47). We calculate $\Xi(\kappa, m, n)$ for $\kappa = 0, 1$ and $m + n \leq 4$. The results are listed below.

$$\Xi(0,0,0)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) dydh = \sqrt{\frac{2}{\pi t}}$$

$$\Xi(1,0,0)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (2h-y-x)dydh = 1$$

$$\Xi(1,1,0)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (2h-y-x)(h-x)dydh$$
$$= \frac{2t^{1/2}}{\sqrt{2\pi}}$$

$$\Xi(1,2,0)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (2h-y-x)(h-x)^2dydh$$
$$= t$$

$$\Xi(1,3,0)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (2h-y-x)(h-x)^3dydh$$
$$= 2\sqrt{\frac{2}{\pi}}t^{3/2}$$

$$\Xi(1,4,0)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (2h-y-x)(h-x)^4dydh$$
$$= 3t^2$$

$$\Xi(1,0,1)$$
$$= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right) (2h-y-x)(y-x)dydh$$

$$= 0$$

$$\begin{aligned}
&\Xi(1,0,2)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(y-x)^2 dydh\\
&= t
\end{aligned}$$

$$\begin{aligned}
&\Xi(1,0,3)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(y-x)^3 dydh\\
&= 0
\end{aligned}$$

$$\begin{aligned}
&\Xi(1,0,4)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(y-x)^4 dydh\\
&= 3t^2
\end{aligned}$$

$$\begin{aligned}
&\Xi(1,1,3)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(h-x)(y-x)^3 dydh\\
&= \frac{3t^2}{2}
\end{aligned}$$

$$\begin{aligned}
&\Xi(1,3,1)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(h-x)^3(y-x) dydh\\
&= \frac{9t^2}{4}
\end{aligned}$$

$$\begin{aligned}
&\Xi(1,2,2)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(h-x)^2(y-x)^2 dydh\\
&= 2t^2
\end{aligned}$$

$$\begin{aligned}
&\Xi(1,1,1)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(h-x)(y-x) dydh\\
&= \frac{t}{2}
\end{aligned}$$

$$
\begin{aligned}
&\Xi(1,2,1)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(h-x)^2(y-x)dydh\\
&= \frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2}
\end{aligned}
$$

$$
\begin{aligned}
&\Xi(1,1,2)\\
&= \frac{2}{t\sqrt{2\pi t}} \int_x^\infty \int_{-\infty}^h \exp\left(-\frac{2(h-y)(h-x)}{t} - \frac{(y-x)^2}{2t}\right)(2h-y-x)(h-x)(y-x)^2dydh\\
&= \frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2}
\end{aligned}
$$

10.4.2 Explicit calculation of the coefficients in the expansion

In this paragraph we prove the missing facts for Theorem 8.3.3.1. This means, we explicitly calculate the coefficients in formula (8.83). We make use of the table of integrals in the previous paragraph, see Appendix 10.4.1. But first note that the Taylor polynomial of degree 4 at $(x,x) \in \mathbb{R}^2$ for a function $g \in \mathcal{C}^4(\mathbb{R}^2,\mathbb{R})$ is given by

$$
\begin{aligned}
&T_4 g((x,x),(h,y))\\
&= g(x,x) + \sum_{\substack{m,n=0\\ m+n\geq 1}}^{4} d_{m,n}\, g_{(m,n)}(x,x)(h-x)^m(y-x)^n\\
&= g(x,x)\\
&\quad + g_{(1,0)}(x,x)(h-x) + g_{(0,1)}(x,x)(y-x)\\
&\quad + \frac{1}{2}g_{(2,0)}(x,x)(h-x)^2 + \frac{1}{2}g_{(0,2)}(x,x)(y-x)^2 + g_{(1,1)}(x,x)(h-x)(y-x)\\
&\quad + \frac{1}{6}g_{(3,0)}(x,x)(h-x)^3 + \frac{1}{6}g_{(0,3)}(x,x)(y-x)^3 + \frac{1}{2}g_{(2,1)}(x,x)(h-x)^2(y-x)\\
&\qquad + \frac{1}{2}g_{(1,2)}(x,x)(h-x)(y-x)^2\\
&\quad + \frac{1}{24}g_{(4,0)}(x,x)(h-x)^4 + \frac{1}{24}g_{(0,4)}(x,x)(y-x)^4 + \frac{1}{6}g_{(3,1)}(x,x)(h-x)^3(y-x)\\
&\qquad + \frac{1}{6}g_{(1,3)}(x,x)(h-x)(y-x)^3 + \frac{1}{4}g_{(2,2)}(x,x)(h-x)^2(y-x)^2. \qquad (10.153)
\end{aligned}
$$

By Corollary 8.3.2.7 it remains to calculate

$$
\int_x^\infty \int_{-\infty}^h (h-x)^m(y-x)^n f(t,x,h,y)dydh
$$

$$= \sum_{j=0}^{4-(m+n)} \sum_{k=1}^{j} \frac{c_{j,k}}{\sqrt{t}^k} \Xi(1,m,n+k)\eta_j(t,x) + O(t^{5/2}) \tag{10.154}$$

for the respective values of m and n. The coefficients $c_{j,k}$ are inferred from the j^{th} Hermite polynomial H_j in the following way

$$c_{j,k} = \frac{d^k}{dx^k} H_j(x)\Big|_{x=0}. \tag{10.155}$$

The results are displayed in the sequel.

The coefficients belonging to the first derivatives

We have

$$\begin{aligned}
&\int_x^\infty \int_{-\infty}^h g_{(1,0)}(x,x)(h-x)f(t,x,h,y)dydh \\
&= g_{(1,0)}(x,x)\int_x^\infty \int_{-\infty}^h (h-x)2\exp\left(-\frac{2(h-y)(h-x)}{t}\right)\frac{(2h-y-x)}{t}p(t,x,y)dydh \\
&\quad + O(t^{5/2}) \\
&= g_{(1,0)}(x,x)\frac{2}{t}\int_x^\infty \int_{-\infty}^h (h-x)(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right) \\
&\quad \times \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(y-x)^2}{2t}\right)\sum_{j=0}^{3}\eta_j(t,x)H_j\left(\frac{y-x}{\sqrt{t}}\right)dydh + O(t^{5/2}) \\
&= g_{(1,0)}(x,x)\left\{\Xi(1,1,0)\eta_0(t,x) + \frac{1}{\sqrt{t}}\Xi(1,1,1)\eta_1(t,x)\right. \\
&\quad \left. + \left(\frac{1}{2t}\Xi(1,1,2) - \frac{1}{2}\Xi(1,1,0)\right)\eta_2(t,x) + \underbrace{\left(\frac{1}{6t^{3/2}}\Xi(1,1,3) - \frac{1}{2\sqrt{t}}\Xi(1,1,1)\right)}_{=0}\eta_3(t,x)\right\} \\
&\quad + O(t^{5/2}) \\
&= g_{(1,0)}(x,x)\left\{\sqrt{\frac{2}{\pi}}\sqrt{t} + \frac{1}{2}\mu(x)t + \frac{1}{4}t^2\left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x)\right) + \frac{1}{3}\frac{1}{\sqrt{2\pi}}t^{3/2}\left(\mu'(x)+\mu(x)^2\right)\right\} \\
&\quad + O(t^{5/2})
\end{aligned} \tag{10.156}$$

and

$$\begin{aligned}
&\int_x^\infty \int_{-\infty}^h g_{(0,1)}(x,x)(y-x)f(t,x,h,y)dydh \\
&= g_{(0,1)}(x,x)\int_x^\infty \int_{-\infty}^h (y-x)2\exp\left(-\frac{2(h-y)(h-x)}{t}\right)\frac{(2h-y-x)}{t}p(t,x,y)dydh + O(t^{5/2}) \\
&= g_{(0,1)}(x,x)\frac{2}{t}\int_x^\infty \int_{-\infty}^h (y-x)(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right)
\end{aligned}$$

$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{3} \eta_j(t,x) H_j\left(\frac{y-x}{\sqrt{t}}\right) dydh + O(t^{5/2})$$

$$= g_{(0,1)}(x,x) \left\{ \underbrace{\Xi(1,0,1)}_{=0} \eta_0(t,x) + \frac{1}{\sqrt{t}} \Xi(1,0,2) \eta_1(t,x) \right.$$

$$\left. + \underbrace{\left(\frac{1}{2t}\Xi(1,0,3) - \frac{1}{2}\Xi(1,0,1)\right)}_{=0} \eta_2(t,x) + \underbrace{\left(\frac{1}{6t^{3/2}}\Xi(1,0,4) - \frac{1}{2\sqrt{t}}\Xi(1,0,2)\right)}_{=0} \eta_3(t,x) \right\}$$

$$+ O(t^{5/2})$$

$$= g_{(0,1)}(x,x) \left\{ \mu(x)t + \frac{1}{2}t^2 \left(\mu(x)\mu'(x) + \frac{1}{2}\mu''(x) \right) \right\} + O(t^{5/2}). \tag{10.157}$$

The coefficients belonging to the second derivatives

First, we have

$$\int_x^\infty \int_{-\infty}^h \frac{1}{2} g_{(2,0)}(x,x)(h-x)^2 f(t,x,h,y) dydh$$

$$= g_{(2,0)}(x,x) \frac{1}{t} \int_x^\infty \int_{-\infty}^h (h-x)^2 (2h-y-x) \exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$

$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{2} \eta_j(t,x) H_j\left(\frac{y-x}{\sqrt{t}}\right) dydh + O(t^{5/2})$$

$$= g_{(2,0)}(x,x) \frac{1}{2} \left\{ \Xi(1,2,0)\eta_0(t,x) + \frac{1}{\sqrt{t}} \Xi(1,2,1)\eta_1(t,x) \right.$$

$$\left. + \left(\frac{1}{2t}\Xi(1,2,2) - \frac{1}{2}\Xi(1,2,0)\right) \eta_2(t,x) \right\}$$

$$+ O(t^{5/2})$$

$$= g_{(2,0)}(x,x) \frac{1}{2} \left\{ t + \frac{4}{3}\sqrt{\frac{2}{\pi}} \mu(x) t^{3/2} + \frac{1}{2} t^2 (\mu'(x) + \mu(x)^2) \right\} + O(t^{5/2}). \tag{10.158}$$

Moreover, we have

$$\int_x^\infty \int_{-\infty}^h \frac{1}{2} g_{(0,2)}(x,x)(y-x)^2 f(t,x,h,y) dydh$$

$$= g_{(0,2)}(x,x) \frac{1}{t} \int_x^\infty \int_{-\infty}^h (y-x)^2 (2h-y-x) \exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$

$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{2} \eta_j(t,x) H_j\left(\frac{y-x}{\sqrt{t}}\right) dydh + O(t^{5/2})$$

$$= g_{(0,2)}(x,x) \frac{1}{2} \left\{ \Xi(1,0,2)\eta_0(t,x) + \frac{1}{\sqrt{t}} \Xi(1,0,3)\eta_1(t,x) \right.$$

$$+ \left(\frac{1}{2t}\Xi(1,0,4) - \frac{1}{2}\Xi(1,0,2)\right)\eta_2(t,x)\Bigg\} + O(t^{5/2})$$
$$= g_{(0,2)}(x,x)\frac{1}{2}\left\{t + t^2(\mu'(x) + \mu(x)^2)\right\} + O(t^{5/2}) \tag{10.159}$$

and

$$\int_x^\infty \int_{-\infty}^h \frac{1}{2} g_{(1,1)}(x,x)(h-x)(y-x) f(t,x,h,y)\,dydh$$
$$= g_{(1,1)}(x,x)\frac{2}{t}\int_x^\infty \int_{-\infty}^h (y-x)^2(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$
$$\times \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(y-x)^2}{2t}\right)\sum_{j=0}^{2}\eta_j(t,x)H_j\left(\frac{y-x}{\sqrt{t}}\right)dydh + O(t^{5/2})$$
$$= g_{(1,1)}(x,x)\Bigg\{\Xi(1,1,1)\eta_0(t,x) + \frac{1}{\sqrt{t}}\Xi(1,1,2)\eta_1(t,x)$$
$$+ \left(\frac{1}{2t}\Xi(1,1,3) - \frac{1}{2}\Xi(1,1,1)\right)\eta_2(t,x)\Bigg\} + O(t^{5/2})$$
$$= g_{(1,1)}(x,x)\left\{\frac{1}{2}t + \frac{4}{3}\sqrt{\frac{2}{\pi}}\mu(x)t^{3/2} + \frac{1}{2}t^2(\mu'(x) + \mu(x)^2)\right\} + O(t^{5/2}). \tag{10.160}$$

The coefficients belonging to the third derivatives

We find

$$\int_x^\infty \int_{-\infty}^h \frac{1}{6} g_{(3,0)}(x,x)(h-x)^3 f(t,x,h,y)\,dydh$$
$$= g_{(3,0)}(x,x)\frac{2}{6t}\int_x^\infty \int_{-\infty}^h (h-x)^3(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$
$$\times \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(y-x)^2}{2t}\right)\sum_{j=0}^{1}\eta_j(t,x)H_j\left(\frac{y-x}{\sqrt{t}}\right)dydh + O(t^{5/2})$$
$$= g_{(3,0)}(x,x)\frac{1}{6}\left\{\Xi(1,3,0)\eta_0(t,x) + \frac{1}{\sqrt{t}}\Xi(1,3,1)\eta_1(t,x)\right\} + O(t^{5/2})$$
$$= g_{(3,0)}(x,x)\frac{1}{6}\left\{2\sqrt{\frac{2}{\pi}}t^{3/2} + \frac{9}{4}\mu(x)t^2\right\} + O(t^{5/2}). \tag{10.161}$$

Moreover, we have

$$\int_x^\infty \int_{-\infty}^h \frac{1}{6} g_{(0,3)}(x,x)(y-x)^3 f(t,x,h,y)\,dydh$$
$$= g_{(0,3)}(x,x)\frac{2}{6t}\int_x^\infty \int_{-\infty}^h (y-x)^3(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$

$$
\begin{aligned}
&\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{1} \eta_j(t,x) H_j\left(\frac{y-x}{\sqrt{t}}\right) dydh + O(t^{5/2}) \\
&= g_{(0,3)}(x,x)\frac{1}{6}\left\{\Xi(1,0,3)\eta_0(t,x) + \frac{1}{\sqrt{t}}\Xi(1,0,4)\eta_1(t,x)\right\} + O(t^{5/2}) \\
&= g_{(0,3)}(x,x)\frac{1}{2}\mu(x)t^2 + O(t^{5/2}) \qquad (10.162)
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_x^\infty \int_{-\infty}^h \frac{1}{2} g_{(2,1)}(x,x)(h-x)^2(y-x) f(t,x,h,y)dydh \\
&= g_{(2,1)}(x,x)\frac{1}{t}\int_x^\infty \int_{-\infty}^h (h-x)^2(y-x)(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right) \\
&\quad \times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{1} \eta_j(t,x) H_j\left(\frac{y-x}{\sqrt{t}}\right) dydh + O(t^{5/2}) \\
&= g_{(2,1)}(x,x)\frac{1}{2}\left\{\Xi(1,2,1)\eta_0(t,x) + \frac{1}{\sqrt{t}}\Xi(1,2,2)\eta_1(t,x)\right\} + O(t^{5/2}) \\
&= g_{(2,1)}(x,x)\frac{1}{2}\left\{\frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2} + 2\mu(x)t^2\right\} + O(t^{5/2}). \qquad (10.163)
\end{aligned}
$$

And finally, we have

$$
\begin{aligned}
&\int_x^\infty \int_{-\infty}^h \frac{1}{2} g_{(1,2)}(x,x)(h-x)(y-x)^2 f(t,x,h,y)dydh \\
&= g_{(1,2)}(x,x)\frac{1}{t}\int_x^\infty \int_{-\infty}^h (h-x)(y-x)^2(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right) \\
&\quad \times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \sum_{j=0}^{1} \eta_j(t,x) H_j\left(\frac{y-x}{\sqrt{t}}\right) dydh + O(t^{5/2}) \\
&= g_{(1,2)}(x,x)\frac{1}{2}\left\{\Xi(1,1,2)\eta_0(t,x) + \frac{1}{\sqrt{t}}\Xi(1,1,3)\eta_1(t,x)\right\} + O(t^{5/2}) \\
&= g_{(1,2)}(x,x)\frac{1}{2}\left\{\frac{4}{3}\sqrt{\frac{2}{\pi}}t^{3/2} + \frac{3}{2}\mu(x)t^2\right\} + O(t^{5/2}). \qquad (10.164)
\end{aligned}
$$

The coefficients belonging to the fourth derivatives

We find that

$$
\begin{aligned}
&\int_x^\infty \int_{-\infty}^h \frac{1}{24} g_{(4,0)}(x,x)(h-x)^4 f(t,x,h,y)dydh \\
&= g_{(4,0)}(x,x)\frac{2}{24t}\int_x^\infty \int_{-\infty}^h (h-x)^4(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right)
\end{aligned}
$$

$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \underbrace{\eta_0(t,x) H_0\left(\frac{y-x}{\sqrt{t}}\right)}_{=1} dydh + O(t^{5/2})$$
$$= g_{(4,0)}(x,x) \frac{1}{24} \Xi(1,4,0) \underbrace{\eta_0(t,x)}_{=1} + O(t^{5/2})$$
$$= g_{(4,0)}(x,x) \frac{3}{24} t^2 + O(t^{5/2}). \tag{10.165}$$

and

$$\int_x^\infty \int_{-\infty}^h \frac{1}{24} g_{(0,4)}(x,x)(y-x)^4 f(t,x,h,y) dydh$$
$$= g_{(0,4)}(x,x) \frac{2}{24t} \int_x^\infty \int_{-\infty}^h (y-x)^4 (2h-y-x) \exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$
$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) dydh + O(t^{5/2})$$
$$= g_{(0,4)}(x,x) \frac{1}{24} \Xi(1,0,4) + O(t^{5/2})$$
$$= g_{(0,4)}(x,x) \frac{3}{24} t^2 + O(t^{5/2}). \tag{10.166}$$

Moreover, we infer that

$$\int_x^\infty \int_{-\infty}^h \frac{1}{6} g_{(3,1)}(x,x)(h-x)^3 (y-x) f(t,x,h,y) dydh$$
$$= g_{(3,1)}(x,x) \frac{2}{6t} \int_x^\infty \int_{-\infty}^h (h-x)^3 (y-x)(2h-y-x) \exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$
$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) dydh + O(t^{5/2})$$
$$= g_{(3,1)}(x,x) \frac{1}{6} \Xi(1,3,1) + O(t^{5/2})$$
$$= g_{(3,1)}(x,x) \frac{1}{6} \frac{9}{4} t^2 + O(t^{5/2}) \tag{10.167}$$

and

$$\int_x^\infty \int_{-\infty}^h \frac{1}{6} g_{(1,3)}(x,x)(h-x)(y-x)^3 f(t,x,h,y) dydh$$
$$= g_{(1,3)}(x,x) \frac{2}{6t} \int_x^\infty \int_{-\infty}^h (h-x)(y-x)^3 (2h-y-x) \exp\left(-\frac{2(h-y)(h-x)}{t}\right)$$
$$\times \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) dydh + O(t^{5/2})$$
$$= g_{(1,3)}(x,x) \frac{1}{6} \Xi(1,1,3) + O(t^{5/2})$$
$$= g_{(1,3)}(x,x) \frac{1}{6} \frac{3}{2} t^2 + O(t^{5/2}). \tag{10.168}$$

And finally, we have

$$
\begin{aligned}
&\int_x^\infty \int_{-\infty}^h \frac{1}{4} g_{(1,3)}(x,x)(h-x)^2(y-x)^2 f(t,x,h,y)\,dydh \\
&= g_{(2,2)}(x,x)\frac{2}{4t}\int_x^\infty \int_{-\infty}^h (h-x)^2(y-x)^2(2h-y-x)\exp\left(-\frac{2(h-y)(h-x)}{t}\right) \\
&\quad \times \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(y-x)^2}{2t}\right) dydh + O(t^{5/2}) \\
&= g_{(2,2)}(x,x)\frac{1}{4}2t^2 + O(t^{5/2}).
\end{aligned} \tag{10.169}
$$

Bibliography

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, 1964.

[2] R.A. Adams. *Sobolev Spaces.* Academic Press, 1978.

[3] Y. Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Annals of Statistics*, 1:223–262, 2002.

[4] Y. Aït-Sahalia. Closed-form likelihood expansions for multivariate diffusions. *Annals of Statistics*, 36:906–937, 2008.

[5] H.W. Alt. *Lineare Funktionalanalysis.* Springer, 4th edition, 2002.

[6] P. Baldi and L. Caramellino. Asymptotics of hitting probabilities for general one-dimensional pinned diffusions. *Annals of Applied Probability*, 12(3):1071–1095, 2002.

[7] R. F. Bass. *Diffusions and Elliptic Operators.* Springer, 1998.

[8] A.T. Bharucha-Reid. *Elements of the Theory of Markov Processes and Their Applications.* Dover, 1960.

[9] B. M. Bibby, M. Jacobsen, and M. Sørensen. Estimating functions for discretely sampled diffusion-type models. *Y. Aït-Sahalia & L. P. Hansen, editors, Handbook of Financial Econometrics. Elsevier. To appear.*, 2009.

[10] B.M. Bibby. Optimal combination of martingale estimating functions for discretely observed diffusion processes. *Preprint, Department of Mathematical Sciences, University of Copenhagen*, 1994.

[11] B.M. Bibby and M. Sørensen. Martingale estimating functions for discretely observed diffusion processes. *Bernoulli*, 1(1/2):17–39, 1995.

[12] B.M. Bibby and M. Sørensen. On estimation for discretely observed diffusions: A review. *Theory of Stochastic Processes*, 2:49–56, 1996.

[13] B.M. Bibby, M. Jacobsen, and M. Sørensen. Estimating functions for discretely sampled diffusion-type models. *Preprint*, 2004. URL `http://www.math.ku.dk/~michael`.

[14] P. Billingsley. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, 12:788–792, 1961.

[15] P. Billingsley. *Convergence of Probability Measures.* Wiley, 2nd edition, 1999.

[16] A. Borodin and P. Salminen. *Handbook of Brownian Motion - Facts and Formulae.* Birkhäuser, 2nd edition, 2002.

[17] K. Borovkov and A. N. Downes. First passage densities and crossing probabilities for diffusion processes. *Methodol. Comput. Appl. Probab.*, 10(4):621–644, 2008.

[18] D.A. Darling and A.J.E. Siegert. The first passage problem for a continuous Markov process. *Annals of Mathematical Statistics*, 24:624–639, 1953.

[19] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Barlett Publishers, Boston., 1993.

[20] M. Dominé. First passage time distribution of a Wiener process with drift concerning two elastic barriers. *Journal of Applied Probability*, 33:164–175, 1996.

[21] S.D. Eidelman and N.V. Zhitarashu. *Parabolic Boundary Value Problems.* Birkhäuser, 1998. Operator Theory, Advances and Applications, Vol. 101.

[22] L. Elie. Equivalent de la densité d'une diffusion en temps petits cas des point proches. *Astérisque*, 84-85:55–71, 1980.

[23] L.C. Evans. *Partial Differential Equations.* AMS, 1997.

[24] P. Fitzsimmons, J. Pitman, and M. Yor. Markovian bridges: construction, palm interpretation, and splicing. *In Seminar on Stochastic Processes*, 1992.

[25] W.H. Fleming and M.R. James. Asymptotic series and exit time probabilities. *The Annals of Probability*, 20:1369–1384, 1992.

[26] D. Florens-Zmirou. Approximate discrete time schemes for statistics of diffusion processes. *Statistics*, 20:547–557, 1989.

[27] O. Forster. *Analysis 3.* Vieweg, 3rd edition, 1984.

[28] A. Friedman. *Stochastic Differential Equations and Applications, Volume 1.* Academic Press, New York, 1975.

[29] A. Friedman. *Stochastic Differential Equations and Applications, Volume 1.* Academic Press, New York, 1976.

[30] M.B. Garman and M.J. Klass. On the estimation of security price volatility from historical data. *Journal of Business*, 53(1):67–78, 1980.

[31] V. Genon-Catalot, T. Jeantheau, and C. Laredo. Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli*, 6(6):1051–1079, 2000.

[32] V.P. Godambe and C.C. Heyde. Quasi likelihood and optimal estimation. *International Statistical Review*, 55:231–244, 1987.

[33] P. Hall and C.C. Heyde. *Martingale Limit Theory and its Applications.* Academic Press, New York, 1980.

[34] C.C. Heyde. Fixed sample and asymptotic optimality for classes of estimating functions. *Contemporary Mathematics*, 80:241–247, 1988.

[35] P. Howard and K. Zumbrun. Shift invariance of the occupation time of the Brownian bridge process. *Statistics and probability letters*, 45(4):379–382, 1999.

[36] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North Holland mathematical library ; 24. North-Holland, 2nd edition, 1989.

[37] M. Jacobsen. Discretely observed diffusions: Classes of estimating functions and small-Δ-optimality. *Scandinavian Journal of Statistics*, 28:123–149, 2001.

[38] M. Jacobsen. Optimality and small-Δ-optimality of martingale estimating functions. *Bernoulli*, 8:643–668, 2002.

[39] J. Jacod and Y. Aït-Sahalia. Is Brownian motion necessary to model high frequency data? *Annals of Statistics*, 38:3093–3128, 2010.

[40] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, 2nd edition, 2003.

[41] O. Kallenberg. *Foundations of Modern Probability.* Springer, 2nd edition, 2002.

[42] J. Kampen. On the WKB-expansion of parabolic equations and applications. *SSRN*, 2006.

[43] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 2nd edition, 1998.

[44] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.*, 24:211–229, 1997.

[45] M. Kessler. Simple and explicit estimating functions for a discretely observed diffusion. *Scand. J. Statist.*, 27:65–82, 2000.

[46] M. Kessler and M. Sørensen. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Research Report No. 332. Department of Theoretical Statistics. University of Aarhus.*, 1995.

[47] O.A. Ladyshenskaja, V.A. Solonnikov, and N.N. Ual'ceva. *Linear and Quasilinear Equations of Parabolic Type.* Amerian Mathematical Society, Providence, RI, 1968.

[48] E.L. Lehmann and G. Casella. *Theory of point estimation.* Springer, 2nd edition, 1998. Operator Theory, Advances and Applications, Vol. 101.

[49] M. Magdon-Ismail and A.F. Atiya. Volatility estimation using high, low, and close data - a maximum likelihood approach. *Computational Finance (CF2000), Proceedings*, 2000.

[50] D. Nualart. *The Malliavin Calculus and Related Topics.* Springer, 2nd edition, 2005.

[51] M. Parkinson. The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65, January 1980.

[52] A. R. Pedersen. Quasi-likelihood inference for discretely observed diffusion processes. *Research Report No. 295, Department of Theoretical Statistics, University of Aarhus*, 1994.

[53] A. R. Pedersen. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics*, 22(1):55–71, 1995.

[54] J. L. Pedersen, L. Alili, and P. Patie. Representations of the first hitting time density of an Ornstein-Uhlenbeck process. *Stochastic Models*, 21:967–980, 2005.

[55] G. Peskir. Limit at zero of the Brownian first-passage density. *Probab. Theory and Relat. Fields*, 124:100–111, 2002.

[56] M. Reiß. Stochastic differential equations. *Lecture Notes*, 2007. URL `http://www.mathematik.hu-berlin.de/~mreiss`.

[57] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, 2nd edition, 1994.

[58] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, 3rd edition, 1999.

[59] L. C. G. Rogers. Volatility estimation with price quanta. *Mathematical Finance*, 8: 277–290(14), July 1998.

[60] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, 2000. Itô calculus, Reprint of the second (1994) edition.

[61] L.C.G. Rogers and S.E. Satchell. Estimating variance form high, low and closing prices. *The Annals of Applied Probability*, 4:504–512, 1991.

[62] L.C.G. Rogers and L. Shepp. The correlation of the maxima of correlated Brownian motions. *Journal of Applied Probability*, 43:880–883, 2006.

[63] L.C.G. Rogers and F. Zhou. Estimating correlation from high, low, opening and closing prices. *Annals of Applied Probability*, 18(2):813–823, 2008.

[64] E. Shargorodsky and A.V. Sobolev. Quasi-conformal mappings and periodic spectral problems in dimension two. *Preprint*, 2001. URL `http://arxiv.org/pdf/math.SP/0109216.pdf`.

[65] M. Sørensen. Estimating functions for discretely sampled diffusion-type models: A review. *IMS Lecture Notes - Monograph Series*, 32:305–325, 1997.

[66] M. Sørensen. On asymptotics of estimating functions. *Brazilian Journal of Probability and Statistics*, 13:111–136, 1999.

[67] M. Sørensen. Efficient estimation for ergodic diffusions sampled at high frequency. *Preprint, Department of Mathematical Sciences, University of Copenhagen*, 2007.

[68] D. W. Stroock. *Partial Differential Equations for Probabilists.* Cambridge, 2008.

[69] D.W. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes.* Springer, 2006.

[70] A.L. Sweet and J.C. Hardin. Solutions for some diffusion processes with two barriers. *Journal of Applied Probability*, 7:423–431, 1970.

[71] E.T. Whittaker and G.N. Watson. *A Course of Modern Analysis.* Cambridge University Press, 1952.

[72] N. Yoshida. Estimation for diffusion processes from discrete observations. *Jour. Mult. Anal.*, 41:220–242, 1990.

[73] L. Zhang. Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli*, 12(6):1019–1043, 2006.

[74] L. Zhang, P. A. Mykland, and Y. Aït-Sahalia. A tale of two time scales: determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.*, 100 (472):1394–1411, 2005.

List of Tables

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 29. März 2010 Hartmuth Henkel